Chapter 1
The Functional Neuroanatomy of Face Processing: Insights from Neuroimaging and Implications for Deep Learning

Kalanit Grill-Spector, Kendrick Kay and Kevin S. Weiner

Abstract Face perception is critical for normal social functioning, and is mediated by a cortical network of regions in the ventral visual stream. Comparative analysis between present deep neural network architectures for biometrics and neural architectures in the human brain is necessary for developing artificial systems with human abilities. Neuroimaging research has advanced our understanding regarding the functional architecture of the human ventral face network. Here, we describe recent neuroimaging findings in three domains: (1) the macro- and microscopic anatomical features of the ventral face network in the human brain, (2) the characteristics of white matter connections, and (3) the basic computations performed by population receptive fields within face-selective regions composing this network. Then, we consider how empirical findings can inform the development of accurate computational deep neural networks for face recognition as well as shed light on computational benefits of specific neural implementational features.

Introduction

Face perception is critical for normal social functioning. For example, faces provide key visual information that we use to discriminate one person from another every single day. Since face perception is ecologically and evolutionarily relevant across species [39, 52, 152, 167] a fundamental question is: *What neuroanatomical and functional features of the human brain contribute to the visual perception and recognition of faces*?

K. Grill-Spector (✉) · K.S. Weiner
Department of Psychology, Stanford University, Stanford, CA, USA
e-mail: kalanit@stanford.edu

K. Grill-Spector
Stanford Neurosciences Institute, Stanford University, Stanford, CA, USA

K. Kay
Department of Radiology, University of Minnesota, Minneapolis, MN, USA

B. Bhanu and A. Kumar (eds.), *Deep Learning for Biometrics*,
Advances in Computer Vision and Pattern Recognition,
DOI 10.1007/978-3-319-61657-5_1

To tackle this question, many influential theories regarding the cognitive neuroscience of face perception [12, 29, 39, 58, 67, 71, 123] have examined how regions within the brain fill the representations proposed by classical theories of face perception in cognitive psychology [14, 82, 154]. This approach has been successful in identifying a network of functional regions in the human occipito-temporal cortex that is specialized for processing faces. Indeed, many influential studies have functionally differentiated both ventral from dorsal components of this network [39], as well as functional regions from one another within these dorsal and ventral components, respectively [77, 110, 112, 125, 129, 165, 169]. This chapter focuses on the ventral component of this network and differs from prior papers and book chapters in two main ways. First, while prior articles focus on the face network as an independent system, a repetitive theme throughout this chapter is that the ventral face network is embedded within the visual system more generally. Thus, we discuss and propose that visual processing in regions outside the face network and their interaction with the face network—for example, through white matter connections—is meaningful and contributes to the efficiency of face processing. Second, in addition to understanding the functional characteristics of each region within the ventral face network, this chapter zooms into the cellular structure of neural tissue composing each region. Both of these topics are important and necessary stepping-stones toward building a mechanistic model that would inform how the anatomical structure of the face network subserves computations underlying fast and efficient face recognition.

From both neuroscience and computational perspectives, a complete mechanistic model explaining the functional neuroanatomy of face perception would (1) define each component of the ventral face network, (2) determine the anatomical features of each component, as well as their connections, (3) understand the functional characteristics (e.g., the representations and information) contained within each component, (4) derive the computations within and across components to the point that they can be modeled and cross-validated, and (5) provide an understanding regarding how anatomical features of the underlying neural circuits and their connections implement computations relevant for face perception and recognition.

This chapter synthesizes recent findings and shows that the field has made significant progress toward generating this model. First, we describe functional characteristics of the human ventral face network from well-established research findings. Second, we summarize novel macro- and microanatomical features of the ventral face network as well as features of white matter connections. Third, we discuss basic computations of the ventral face network performed by population receptive fields (pRFs). In the fourth section, we consider how recent empirical findings regarding the human ventral face network could be implemented and tested in computational models including deep convolutional neural networks (CNNs) that contain architectural features inspired by the hierarchical architecture of the ventral stream [41, 121, 136]. While deep CNN architectures are broadly "neurally inspired" by neurobiological architectural features of the ventral visual stream, they also differ from the neural architecture in the human brain in several fundamental ways (Table 1.1). Therefore, in this fourth section, we highlight some of architectural and functional features that have not yet been implemented in deep neural network architectures

and consider potential computational benefits of specific functional and anatomical features of the neurobiological implementation.

1.1 The Functional Characteristics and Organization of the Ventral Face Network in the Human Brain

1.1.1 Functional Characteristics of the Ventral Face Network

Face-selective regions, which exhibit higher responses to faces compared to other stimuli have been identified with neuroimaging methods first with Positron Emission Tomography [134, 135], then with intracranial electroencephalography [1–3, 95, 116], and later with functional magnetic resonance imaging (fMRI) [72, 94, 108, 114, 150, 165]. Based on this functional characteristic, scientists identify a constellation of face-selective regions using fMRI [58]. In the occipital and temporal lobes, scientists identify face-selective regions in both ventral occipito-temporal cortex (Fig. 1.1a), as well as superior temporal cortex [39, 167]. The former are associated with face perception and recognition [35, 50, 97, 150] and the latter are associated with dynamic aspects of face perception [4, 17, 18, 110, 115, 175]. As the focus of this chapter is understanding the neural basis of face recognition, we focus on three regions of the ventral face network: IOG-faces, pFus-faces, and mFus-faces[1] (Fig. 1.1a). The former region is synonymous with the occipital face area (OFA, [42, 109]). The latter two regions are anatomically and functionally distinct components of the fusiform face area (FFA, [72, 165, 166]): pFus-faces is synonymous with FFA-1 [108] while mFus-faces is synonymous with FFA-2 [108]. Additional face-selective regions have been identified in the anterior temporal lobe [10, 70, 118, 153], but these regions are not considered part of the core face network (but see [21]) as they are not just driven by visual stimulation and are more elusive due to lower signals and susceptibility artifacts in fMRI.

The basic functional characteristic of functional regions within the ventral face network is higher neural responses to the visual presentation of faces compared to a variety of other stimuli including animate stimuli (such as limbs, bodies, and animals), familiar and unfamiliar objects, scenes, characters, and textures (Fig. 1.1b). Within each region, functional responses to face exemplars are higher than exemplars of other categories [26, 66, 98, 113] (Fig. 1.1c).

This characteristic response is maintained across sessions [13, 106, 165], tasks [15, 165], and stimulus formats (Fig. 1.1d), including photographs [65, 72], line drawings [65, 72], two-tone stimuli [24, 151], texture [36], and spatial frequency [160].

While the preferential response to faces over other stimuli is maintained across image transformations and face information can be read out from distributed

[1] See Appendix for abbreviations and definitions.

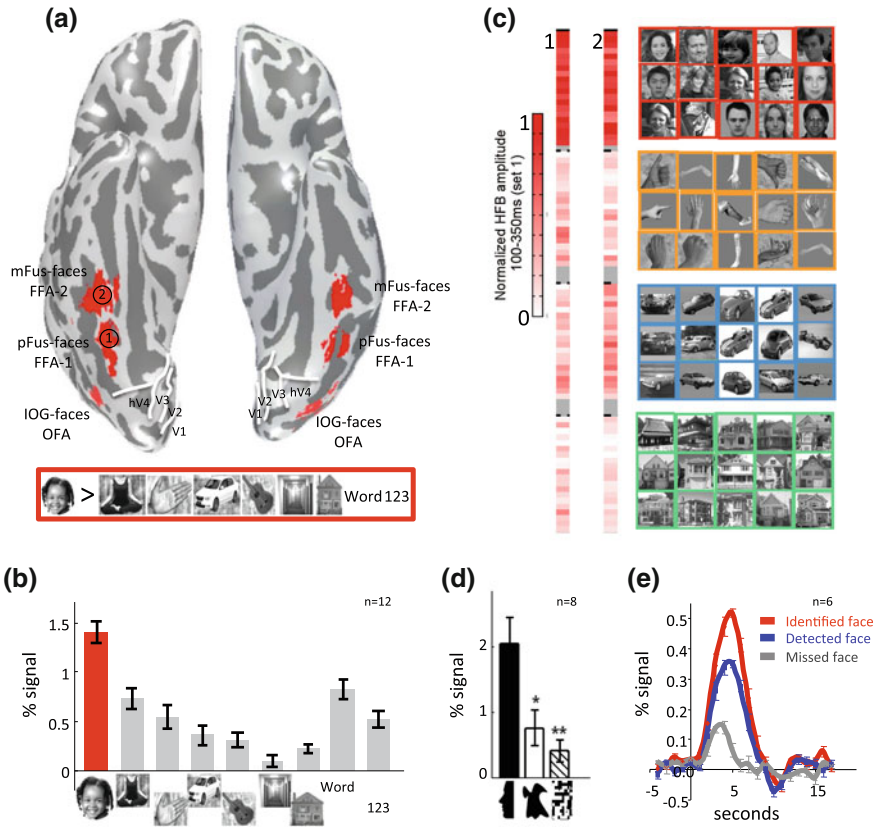

Fig. 1.1 Face-selective regions in human ventral occipito-temporal cortex. **a** Face-selective regions are identified based on higher responses to faces compared to a variety of other stimuli (faces > bodies, objects, places, and characters, t > 3, voxel level). The figure shows an inflated cortical surface of an individual subject depicting the three typical clusters of face-selective regions in ventral occipito-temporal cortex. One cluster is on the inferior occipital gyrus referred to as IOG-faces (also as occipital face area, OFA); a second cluster is on the posterior aspect of the fusiform gyrus, extending to the occipito-temporal sulcus, referred to as pFus-faces (also fusiform face area one, FFA-1); a third cluster is located about 1–1.5 cm more anterior on the lateral aspect of the fusiform gyrus overlapping the anterior tip of the mid-fusiform sulcus (MFS) and is referred to as mFus-faces (also FFA-2). *White lines* boundaries of retinotopic areas. **b** Independent analysis of response amplitudes of mFus-faces showing the typical higher responses to faces compared to other stimuli. Adapted from Stigliani et al. 2015. **c** Responses to single images in pFus- and mFus-faces. Each cell shows the normalized electrocorticography responses to single images averaged over 5 presentations and a 100–350 ms time window. The *first column* shows responses in an intracranially implanted electrode over pFus-faces/FFA-1 and the *second* shows responses from an electrode over mFus-faces/FFA-2. Responses to face images are consistently higher than responses to any of the nonface images. Adapted from [66]. **d** Responses in ventral face-selective regions to face silhouettes are significantly higher than two-tone shapes and scrambled images. Adapted from [24]. **e** Responses in ventral face-selective regions are highest when faces are identified, intermediate when they are detected but not identified, and lowest when they are missed. Adapted from [50]

responses patterns [5, 9, 19, 79, 102, 132], functional responses of the ventral face network are modulated by stimulus position [89, 183], size [183], illumination [49], contrast [126, 183], and viewpoint [49, 79, 102, 159]. For example, responses in ventral face-selective regions are higher for upright than upside down faces [24, 73, 182] and are higher for centrally presented than peripheral faces [57, 89]. Additionally, responses in face-selective regions are modulated by top-down effects, including attention [23, 25, 104], expectation [32, 143], and familiarity [6, 34, 56, 101, 164].

Critically, fMRI-adaptation [47, 49, 51] experiments have been pivotal in showing that responses in ventral face-selective regions are sensitive to face identity. For example, repeating the same face produces reduced responses due to neural adaptation [47, 49, 51, 124] and parametrically increasing the dissimilarity among face identities systematically increases the responses in face-selective regions due to release from adaptation [24, 43, 67, 90, 100]. Additionally, and consistent with behavioral aspects of face perception, (a) neural sensitivity to face identity is higher for upright than upside down faces [44] and (b) both changes in facial features and in the metric relation among features [129, 181] cause a recovery from adaptation which is associated with the perceived change in identity [26, 126, 129, 130]. Finally, neural responses to faces in ventral face-selective regions are correlated with the perception of individual participants [35, 50, 97, 150] and also causally involved in the perception of faces [68, 69, 95, 105, 111, 112, 119, 123]. For example, neural responses within mFus- and pFus-faces are low when faces are present but not detected, intermediate when faces are detected but not identified, and highest when they are identified ([50]; see Fig. 1.1e).

Altogether, these foundational studies reveal that the amplitudes of neural responses in the ventral face network are both higher for faces than nonfaces across formats and correlate with the perception of faces.

1.2 The Neural Architecture and Connections of the Ventral Face Network

1.2.1 The Functional Organization of the Face Network Is Consistent Across Participants

A striking characteristic feature of the functional architecture of the ventral face network is that the cortical location of functional regions is highly consistent across people. At the centimeter scale, face-selective regions are identifiable on specific gyri: occipital face-selective regions are located on the inferior occipital gyrus (IOG, Fig. 1.1a), while face-selective regions in ventral temporal cortex (VTC) are located on the lateral aspect of the fusiform gyrus (FG).

It is important to emphasize that gyri are not small—they are typically several centimeters long and wide and thus, have a rather large surface area. Thus, limiting a functional region to *anywhere* on these macroanatomical structures results in

extensive across-subject variability with low predictability for identifying these regions from cortical folding alone [40]. However, in the last 5 years, we have gained new insights regarding the structural-functional coupling of face-selective regions at the millimeter scale. These advancements have been possible due to rigorous measurements of the variability of the cortical folding patterns of the FG and neighboring sulci [168, 170] as well as precise measurements of the relationship between functional regions and macroanatomical landmarks [99, 165, 170].

One such macroanatomical landmark to note is the mid-fusiform sulcus (MFS), which is a shallow longitudinal sulcus that bisects the fusiform gyrus (FG) into a lateral and medial portion (Fig. 1.2a). While there is anatomical variability across individuals in the length and fractionation of the MFS [170], it serves as a consistent landmark identifying functional representations. For example, irrespective of inter-individual variability in MFS morphology, the anterior tip of the MFS predicts the location of the mid-fusiform face-selective region, identifying about 80% of voxels within mFus-faces/FFA-2 (Fig. 1.2a-left [165, 170]). By comparison, the posterior tip of the MFS predicts about 50% of pFus-faces/FFA-1 (Fig. 1.2a-left). This lower functional-macroanatomical coupling is due to higher anatomical variability of the posterior compared to anterior end of the MFS. Interestingly, the structural-functional coupling extends to large-scale maps spanning several centimeters in VTC. For example, the MFS also identifies a transition within a large-scale animacy map spanning VTC [48] in which voxels that prefer animate stimuli are located lateral to the MFS and voxels that prefer inanimate stimuli are located medial to the MFS (Fig. 1.2a-center). Consequently, this consistent functional-macroanatomical coupling generates a consistent spatial relationship between multiple representations in VTC. That is, face-selective regions are consistently embedded within a larger scale representation of animate stimuli [22, 48, 93].

Given that it is more neurobiologically costly for the brain to spend the energy to generate orderly compared to disorderly representations (and not all visual representations align with anatomical axes on the cortical sheet as described above), these findings raise the following questions: (1) Are there anatomical constraints that contribute to the consistency of this functional architecture? (2) Is there a computational benefit to the regular spatial topography of functional representations in VTC?

1.2.2 The Cytoarchitecture of Face-Selective Regions

What might explain the predictable spatial arrangement of both fine-scale clusters and large-scale functional representations relative to macroanatomical landmarks? Recent evidence indicates that anatomical constraints may underlie the predictable topologies in the VTC.

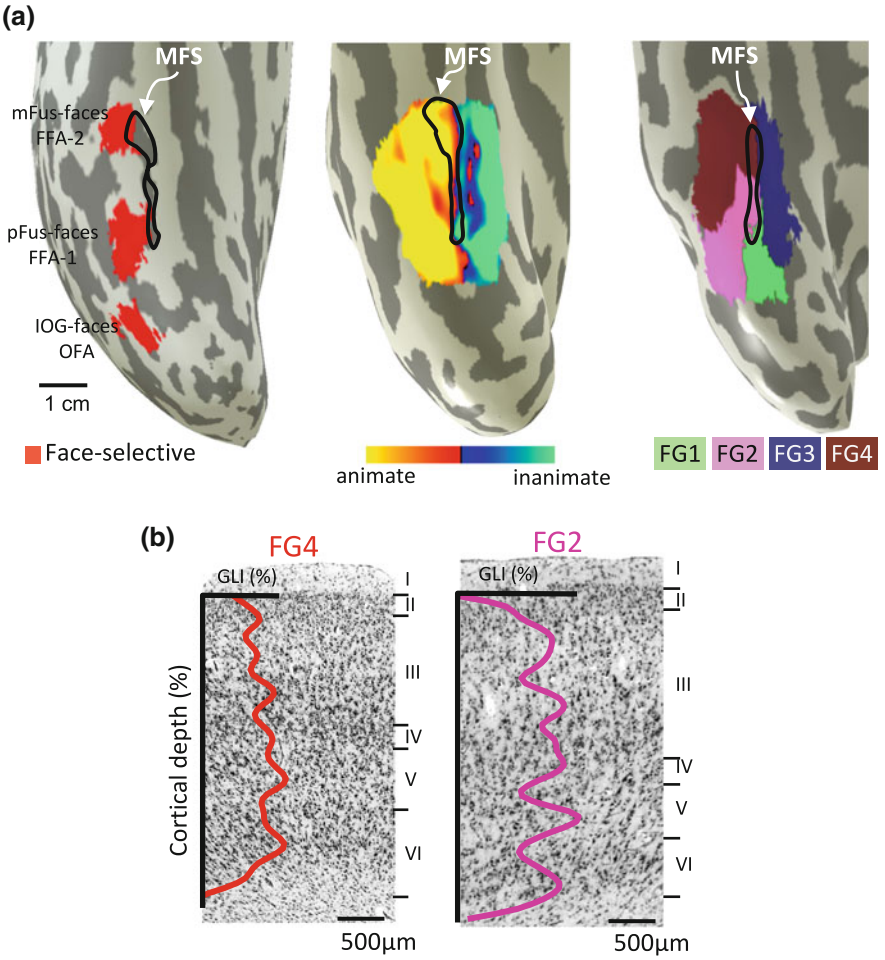

Fig. 1.2 Regular spatial structure of functional and anatomical parcellations of ventral temporal cortex. **a** Multiple representations are aligned to the mid-fusiform sulcus (MFS). Data are shown on an inflated cortical surface zoomed on ventral temporal cortex (VTC) of the right hemisphere of a representative brain. MFS is indicated by the *white outline*. *Left* Face-selective regions (faces > other categories; t > 3, voxel level, *red*) mFus-face/FFA-2 and pFus-faces/FFA-1 are predicted by the anterior and posterior tips of the MFS, respectively. *Center* MFS serves as a boundary between distributed representations of animate and inanimate representations. *Right* The MFS serves as a boundary separating lateral cytoarchitectonic regions FG4 and FG2 from medial cytoarchitectonic regions FG3 and FG1, respectively. Cytoarchitectonic areas are indicated with separate *colors* (see legend). *FG*: fusiform gyrus. **b** Histological slice showing the cell body staining and the gray level index (GLI, *line*) across cortical layers from a representative 20-micron slice through FG4 (*left*) and FG2 (*right*). There are different cell distributions and cell sizes across cortical layers between the two example slices

1.2.2.1 The Cytoarchitecture of the FG

The size, shape, and organization of cells across the six-layered cortical ribbon (the combination of which is referred to as cytoarchitecture) is a well-established criterion to parcellate the brain into areas because differences in cellular structure across cortical layers are believed to be indicative of specialized neural hardware that is utilized for particular brain functions.

In the last 4 years, four cytoarchitectonic areas have been discovered in VTC using data-driven, observer-independent techniques, which are blind to cortical folding. In the posterior aspect of the fusiform gyrus (FG) and neighboring sulci, there are two cytoarchitectonic areas referred to as FG1 and FG2 [20]. FG1 is located on the medial aspect of the FG and is characterized by a columnar structure, while FG2 is situated on the lateral aspect of the FG and has a higher cell density than FG1. Anteriorly, there are two additional cytoarchitectonic areas referred to as FG3 and FG4 ([92]; Fig. 1.2a-right). Many architectural features differentiate FG3 from FG4. For example, FG3 is characterized by a compact and dense layer II, while FG4 is characterized by a less dense layer II compared to FG3, as well as large layers III and V. Interestingly, the MFS not only serves as a landmark identifying face-selective regions and functional transitions in large-scale maps, but also serves as a landmark identifying microarchitectural boundaries separating medial fusiform cytoarchitectonic areas (FG1/FG3) from lateral fusiform cytoarchitectonic areas (FG2/FG4, Fig. 1.2a-right). Specifically, the cytoarchitectonic transition between FG1, medially, to FG2, laterally, occurs 5.41 ± 1.6 mm from the posterior MFS, while the cytoarchitectonic transition between FG3, medially, and FG4, laterally, occurs 1.42 ± .54 mm from the anterior MFS (Fig. 1.2a-right). Since the MFS predicts both functional and anatomical transitions in the human brain, it is natural to ask: Is there relationship between functional regions and cytoarchitectonic areas in the fusiform gyrus?

1.2.2.2 The Relationship Between FG Cytoarchitectonic Areas and the Ventral Face Network

Quantifying the relationship between cytoarchitectonic areas and face-selective regions is challenging because cytoarchitectonic areas are delineated in postmortem brains, while face-selective regions are defined in living brains. Presently, it is impossible to relate these cortical divisions within the same individual. Nevertheless, it is possible to quantitatively relate these structures by aligning them to a common cortical reference frame. This is done using cortex-based alignment, which leverages cortical folding patterns to align one brain to another, irrespective if the brains are from living or postmortem individuals [38, 122]. Implementing this novel approach revealed that functionally defined face-selective regions within the FG are cytoarchitectonically dissociable. That is, different face-selective regions are located within different cytoarchitectonic areas: mFus-faces/FFA2 is largely within FG4 (81% ± 24%, mean ± standard deviation), while pFus-faces/FFA1 is largely within FG2 (49.5% ± 24%) and not in other cytoarchitectonic areas of the FG [173]. These

results suggest that microanatomical properties contribute to the macroanatomical positioning of mFus-faces/FFA-2 and pFus-faces/FFA-1 (even though they are both located on the fusiform gyrus). For example, pFus-faces/FFA-1 displays features of FG2 (Fig. 1.2b), which has a conspicuous layer III with larger pyramidal cells than those of mFus-faces/FFA-2, as well as a prominent and dense layer IV compared to mFus-faces/FFA-2, which has a thin and moderately dense layer IV.

These results have three important theoretical ramifications. First, distinct cytoarchitecture is evidence for differential neural hardware optimized for specialized computations. Thus, it is likely that the cytoarchitectonic differences between mFus-faces/FFA-2 and pFus-faces/FFA-1 are reflective of different computations implemented by these regions. Second, as cytoarchitectonic differences are used to parcellate brain areas, our data suggest that distinct functional regions corresponding to pFus- and mFus-faces, respectively, have distinct cytoarchitectonic structure. Third, since IOG-faces is located outside the FG, it does not overlap with any of the FG cytoarchitectonic areas. This suggests that IOG-faces/OFA is also cytoarchitectonically distinct from pFus-faces/FFA-1 and mFus-faces/FFA-2.

Together these findings suggest that the brain may have different neural hardware for specific computations implemented in each of the three faces-selective regions of the ventral face network. Nevertheless, future research is necessary to elucidate the computations that are produced by this elaborate microcircuitry as well as detailed properties of this circuitry including the precise cell types, their connections, and their 3D structure.

1.2.3 White Matter Connections of the Ventral Face Network

In addition to local cytoarchitecture, there is consensus that white matter connections also constrain the function of the brain [147, 158, 184]. Recent evidence has begun to elucidate the nature of white matter connections of the face network with four main findings. First, ventral face-selective regions IOG-, pFus-, and mFus-faces are highly interconnected with direct white matter connections [54, 117, 171]. Second, longitudinal white matter tracts connect early visual retinotopic areas located outside the face network to ventral face-selective regions [54, 80, 171]. Third, vertical white matter tracts connect dorsal stream visual regions located outside the face network to ventral face-selective regions. For example, portions of the vertical occipital fasciculus (VOF; [146, 172, 179]) connect a retinotopic region in the posterior intraparietal sulcus (IPS-0) and pFus-faces [171]. Fourth, there are distinct white matter tracts associated with the ventral face network compared to networks associated with processing other domains. For example, long-range white matter connections of pFus- and mFus-faces are distinct from white matter connections of a place-selective region in the collateral sulcus (CoS-places/PPA; [45, 117, 128, 149]).

A schematic summarizing these white matter connections is shown in Fig. 1.3. We note that this diagram is incomplete because it does not provide information regarding (1) the entire connectivity of the ventral face network, (2) the direction of

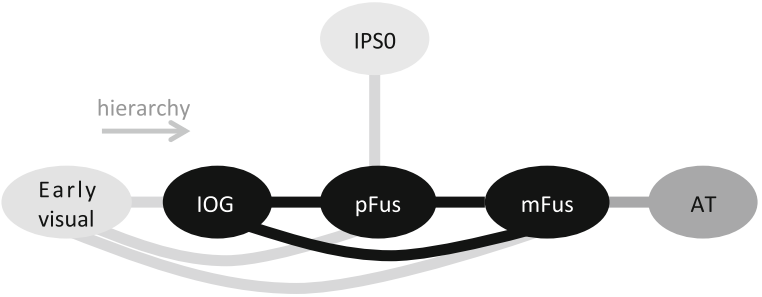

Fig. 1.3 Schematic diagram of white matter tracts of the ventral face network. *Black ovals* indicate core face-selective regions and *gray ovals* indicate regions that are considered external to the core face network. The schematic is arranged such that the hierarchical axis is from left to right. *Acronyms*: face-selective regions: *IOG*: inferior occipital gyrus; *pFus*: posterior fusiform; *mFus*: mid-fusiform: *AT*: anterior temporal; *IPS0*: a region in the intraparietal sulcus that is part of the attention network

connections (though research in animals suggests that they are bidirectional [37]), or (3) the functionality of these white matter tracts.

Nevertheless, the schematic diagram provides four important insights regarding the connectivity features of the ventral face network. First, these findings illustrate hierarchical connections from IOG- to pFus-faces and from pFus- to mFus-faces, which may provide the connectivity scaffolding for hierarchical features of the ventral face network described in the following section. Second, the network contains redundant connections: there are multiple white matter tracts reaching each region. Third, there are bypass connections that do not follow a strict hierarchical organization. For example, mFus-faces, which is considered a later stage of the processing hierarchy following pFus-faces, is connected not only to the preceding stage (pFus-faces), but also to IOG-faces and early visual cortex. Finally, there are vertical white matter tracts connecting IPS-0, which is thought to be part of the attention network [64, 137, 138, 144], to pFus-faces. These vertical tracts may facilitate top-down processing [74].

One important functional feature of redundant and bypass connections is that they may provide network resiliency in the face of injury or disease. For example, recently we had the unique opportunity to measure the face network before (1 month and 4 days) and after (1 and 8 months) a surgical resection of the IOG in a patient (SP) who underwent surgery to treat intractable epilepsy [171]. Surprisingly, downstream regions remained functionally intact despite the resection of the IOG, which would not have been predicted by a strict hierarchical organization [58]. Interestingly, this resiliency of FG face-selective regions is also reported in patients with long-term (>10 years) damage to the inferior occipital cortex [129, 140, 142]. By measuring white matter connections in SP, we identified the longitudinal and vertical tracts discussed above suggesting that these tracts may contribute to the resiliency of the face network after resection by enabling signals to reach these downstream regions using alternate routes from early visual cortex and/or parietal cortex [171].

While identifying white matter tracts of the ventral face network is a major stepping-stone, we recognize that future work is necessary to uncover many remaining unknowns regarding the connectivity of the ventral face network including: (1) What is the functional utility of white matter connections both within the face network as well as to regions outside the network? (2) What is the contribution of white matter connection to the spatial segregation of face-selective regions of the ventral face network? (3) What is the contribution of white matter connections to behavior?

1.3 Computations by Population Receptive Fields in the Ventral Face Network

As described in Sect. 1.2 above, the field has accrued a considerable body of knowledge regarding the functional characteristics of ventral face-selective regions, the anatomical composition and connectivity of regions within the ventral face network, and their role in perception. However, how underlying features contribute to the computations of each region and the network as a whole remains elusive. In this section, we describe progress in understanding basic computations performed across the ventral stream by population receptive fields (pRFs).

A logical starting point for developing computational models of the ventral face network is to determine *receptive field* properties (the region of the visual field within which a stimulus elicits a response from a neuron) for neurons in ventral face-selective regions for three reasons. First, receptive fields are a fundamental aspect of the processing performed by neurons in the visual system [53, 63]. Since neurons with similar RFs are spatially clustered and fMRI measures the population response of all neurons within each brain voxel (volume pixel), we can measure the *population receptive field (pRF)*—the region of the visual field that drives the population of neurons within a voxel [31, 76, 77, 162]. Second, face recognition is thought to require spatial integration across facial features rather than processing isolated facial features [148, 155, 156]. Thus, determining the location and size of pRFs in the ventral face network may inform our understanding of which parts of the face are processed in different stages of the face network. Third, understanding pRFs may shed light on fixation/viewing behavior. For example, when asked to recognize a face, participants typically fixate on the center of the face and the eyes, but when asked to judge the emotion of the face, people also fixate on the mouth [107]. These fixation behaviors suggest that the spatial capacity of face processing is not only limited, but may also be task dependent.

1.3.1 pRF Measurements Reveal a Hierarchical Organization of the Face Network

Recently, we performed a series of experiments in which we built *encoding models* (computational models that explicitly identify a set of features and computations that predict evoked brain responses) that characterize pRFs in the ventral face network. Subjects were scanned while viewing faces at different positions and sizes that systematically tiled the central visual field (12.5°). From these data, we obtained the amplitude of fMRI response of each voxel as a function of face position and size. Then, for each voxel we fit a model that predicts the response by computing the overlap of the stimulus with a 2D Gaussian, followed by a compressive nonlinearity [76, 77]. This model-based analysis (1) provides estimates of pRF properties such as size and eccentricity (distance from fixation), (2) allows comparison of pRF properties across the ventral stream hierarchy, and (3) provides insight into how space is represented in the ventral face network.

PRF mapping shows that voxels in the ventral face network are substantially modulated by the location and spatial extent of faces in the visual field, and our simple pRF model explains these modulations well. PRFs in the ventral face network illustrate several characteristics. First, pRF centers are located in the contralateral visual field [60, 77]—a finding that is consistent with retinotopic organization of visual cortex more generally [62]. That is, pRF centers of voxels in the right hemisphere are centered in the left visual field and vice versa. Second, pRFs in the ventral face network tend to be located close to the center of gaze [77, 176] rather than distributed across the visual field as in early and intermediate retinotopic areas (V1-hV4, Fig. 1.4b). Third, the average pRF size progressively increases from V1 to subsequent retinotopic areas (V2-hV4) and into the face network (IOG-, pFus- and mFus-faces; [77, 176], Fig. 1.4a). This progressive increase of average pRF size is consistent with a hierarchical organization [157]. Fourth, pRF size linearly increases with eccentricity in the face network (Fig. 1.4c; [77, 176] as in earlier regions [162]). Additionally, the slope of this relationship increases across the visual hierarchy. Consequently, the size of pRFs in the ventral face network is larger than their eccentricity. Thus, pRFs in the ventral face network always cover the center of gaze (fovea) and extend into the ipsilateral visual field, processing information from both right and left visual fields.

An intuitive illustration of the visual information processed by pRFs across the ventral processing stream is shown in Fig. 1.4d. Here, we show example pRFs at 1° eccentricity across several regions spanning the ventral stream; these pRFs are superimposed on a face sized to approximate typical conversational distance [91, 96]. The figure illustrates that a V1 pRF processes local information, such as the corner of the eye, while a pRF within hV4 may process an entire facial feature, such as an eye. However, the large and foveal pRFs in the face network process information across multiple facial features from both sides of the face. These data suggest a potential explanation for fixation behavior: when people fixate faces, they attempt to

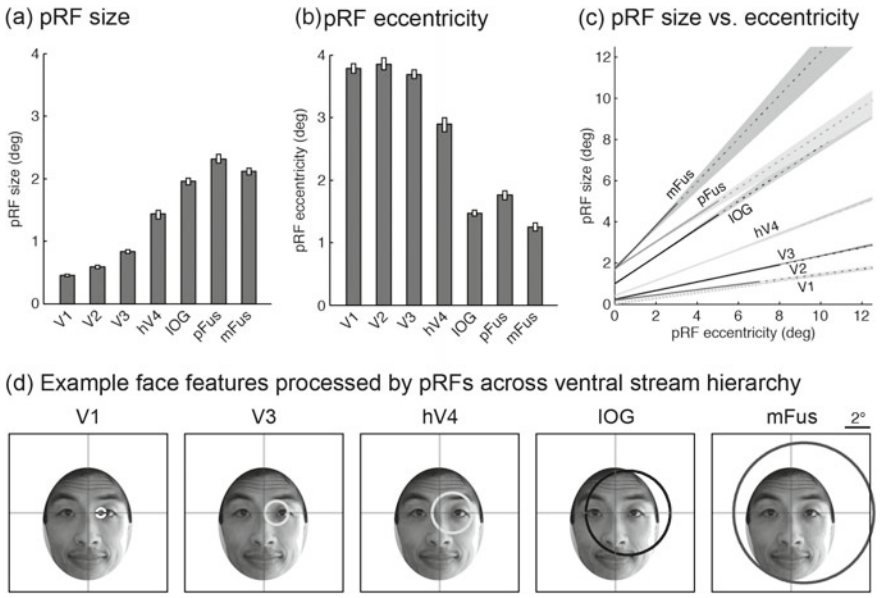

Fig. 1.4 pRFs reveal a hierarchical organization of the ventral face network. **a** Median pRF size for each area from V1 to mFus-faces (*error bars* indicate 68% confidence intervals, CIs). pRF size is defined as the standard deviation of a 2D Gaussian that characterizes the response of the pRF to point stimuli. **b** Median eccentricity of pRF centers for each area. **c** Relationship between pRF size and eccentricity (*shaded area* indicates a 68% CI on a line fitted to the data). **d** Facial features processed across the hierarchy. *Circles* indicate pRFs at 1° eccentricity (as derived from panel **c**). Each *circle* is drawn at +/− 2 pRF sizes. The depicted face is sized to simulate a conversational distance of 1 m (approximately 6.5° based on average male head sizes [91, 96]). Ascending the hierarchy, spatial information is integrated across increasingly large regions of the face, until the latest stages where entire faces are processed by neural populations within a voxel. Adapted from [77, 167]

position pRFs in face-selective regions in order to optimally integrate information across facial features.

Data comparing pRFs in the ventral face network between developmental prosopagnosics (DPs) and typical participants support the idea that spatial integration across facial features obtained by large and central pRFs is necessary for face perception [176]. DPs are individuals without brain damage, but who are impaired at face recognition without showing other visual or cognitive deficits [8, 11, 28, 30, 88]. PRF measurements reveal that pRFs in the face network (and hV4) of DPs are smaller, and rarely extend to the peripheral or ipsilateral visual field compared to typical controls. Notably, across both typicals and DPs, face recognition ability is positively correlated with pRF size in the ventral face network: participants with larger pRFs perform better than those with smaller pRFs ($r(15) = 0.63$, $p<0.007$). In contrast, face recognition ability does not correlate with pRF size in early retinotopic areas. These data provide empirical evidence suggesting that smaller

pRF sizes in DPs may reflect a deficit in spatial integration, consequently affecting face recognition.

1.3.2 Attention Modulates pRF Properties, Enhancing Peripheral Representations Where Visual Acuity Is the Worst

Responses in the ventral face network are not just driven by the stimulus, but are also modulated by internal top-down goals [7, 104, 143, 180]. We characterized how top-down factors modulate pRF properties and spatial information by measuring pRFs under different attentional states [77]. PRFs were estimated separately for different tasks using identical visual stimulation, including a *digit task* during which attention

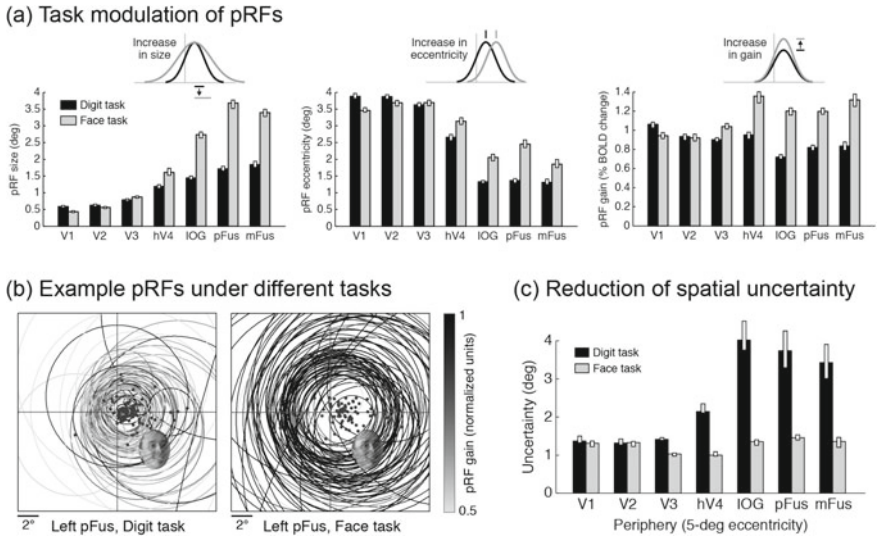

Fig. 1.5 Attention modulates pRF properties in the ventral face network, enhancing spatial representations. PRFs were measured under different tasks using the same stimulus. For the data in this panel, subjects either performed a one-back task on centrally presented digits (digit task) or a one-back task on the presented face (face task) while fixating centrally. **a** Task-induced changes in pRF properties (*bars* indicate median across voxels; *error bars* indicate 68% CIs). pRFs in IOG-, pFus-, and mFus-faces, and hV4 are larger (*left*), more eccentric (*middle*) and have increased gain (*right*) during the face task (*gray*) compared to the digit task (*black*). **b** Tiling of visual field by 100 randomly selected pRFs from *left* pFus-faces (*dots* indicate pRF centers; *circles* indicate pRFs drawn at +/− 2 pRF sizes). An example face is shown at 5-deg eccentricity. **c** Spatial uncertainty in discrimination of stimulus positions. *Bars* indicate amount of uncertainty for reference positions at 5-deg eccentricity (median across angular positions +/− 68% CI). During the digit task, uncertainty in IOG-, pFus-, and mFus-faces is large. However, during the face task, uncertainty is substantially reduced and is commensurate with spatial uncertainty in V1. Adapted from [77, 167]

was directed toward digits presented centrally, at fixation, and a *face task,* during which attention was directed toward faces, which appeared at various locations tiling the visual field (while fixating, subjects performed a one-back judgment on digits and faces, respectively).

Fitting the pRF model separately to brain responses observed under different tasks, we found that pRFs in the ventral face network and hV4 are dependent on the task. In these regions, pRFs are larger (Fig. 1.5a-left), located more peripherally (Fig. 1.5a-center), and have a higher gain (Fig. 1.5a-right) when participants attended to the faces compared to when they attended to digits presented at fixation. In contrast, pRFs in early visual areas V1–V3 were relatively stable and did not substantially change across tasks (Fig. 1.5a).

To obtain an intuitive understanding of the effect of attentional modulation of pRFs in the ventral face network, we visualize the collection of pRFs from left pFus-faces measured in the digit task (Fig. 1.5b-left) and in the face task (Fig. 1.5b-right). As pRFs are larger and more eccentric in the face than digit task, there is extended coverage of the peripheral visual field during the face task compared to the digit task. For example, while a face presented at 5° eccentricity from fixation is processed by only a handful of pRFs during the digit task, it is processed by many pRFs during the face task.

To interpret the change in spatial representations across tasks, we used a model-based *decoding approach* (inferring information from distributed patterns of brain activity across a collection of pRFs) and quantified the spatial uncertainty of the location of a face presented at 5° eccentricity from responses of a collection of pRFs spanning each visual area. Results show that the spatial uncertainly in decoding the location of the face from the collection of pRFs in the face network substantially decreases from the digit to face task (Fig. 1.5c). Surprisingly, the spatial uncertainty in the face network during the face task is similar to that obtained by V1 pRFs which are considerably smaller. These results illuminate another aspect of spatial coding: what determines the spatial resolution of processing by a collection of pRFs is not only their size, but also their scatter. In other words, large and partially overlapping pRFs may provide similar spatial precision as small and nonoverlapping pRFs, consistent with the notion of coarse coding [139, 174].

Our research of pRF properties is only a first stepping-stone of building accurate encoding models of the face network, as we have implemented a rudimentary spatial filter that performs only spatial summation, and the same type of spatial filtering throughout the ventral stream. Thus, important future elaborations of the pRF model would be to (1) include additional dimensions to the model that explain computations related to other aspects of the stimulus (e.g., its shape and/or features), and (2) determine whether and how additional pRF properties vary within each cortical region and across regions. For example, implementing an array of orientation selective filters for each V1 voxel, rather than just Gaussian filters, provides better predictive power in explaining V1 responses to natural images [75]. In the domain of face processing, elaborating pRFs is necessary for explaining basic response properties of face-selective regions not explained by the present pRF model, such as preferential responses and tuning to individual faces [46, 84, 103] and face parts [27, 61].

1.4 Eyes to the Future: Computational Insights from Anatomical and Functional Features of the Face Network

1.4.1 What Is the Computational Utility of the Organized Structure of the Cortical Face Network?

The empirical findings reviewed here reveal an organized and reproducible implementation of a neural processing system for face perception in the human brain. Since generating an organized structure is more effortful than a disorganized structure, it is possible that certain principles reflecting optimized computational strategies are produced from this functional architecture over the course of evolution or development. Therefore, computational insights can be gleaned from the specific features of the physical implementation of the ventral face network in the brain. Here, we highlight some of the architectural and functional features that have not yet been implemented in computational models and consider putative computational aspects of these features that can be tested in future research.

For example, an important aspect of computational models that can be further developed is explicitly modeling how anatomical features may contribute to the formation of dynamic, task-dependent pRFs. The present pRF approach simply treats each task as a distinct entity and estimates a model of the stimulus representation separately for each task. This provides useful insight, but further research is necessary to identify the neural mechanisms that underlie the source of the task modulations originating from other brain areas. To better understand how attention may modulate pRFs, one could consider the characteristics of white matter connections of the face network. For example, top-down connections from cortical regions outside the face network (such as the connection from IPS-0 to pFus-faces) may provide a route for top-down information to flow from parietal regions involved in attentional gating to face-selective regions. Thus, dynamic pRFs may be an outcome of an interaction between a static pRF generated by bottom-up connections and an attention field [81, 120, 141] mediated by top-down connections from IPS-0. Developing new encoding models that incorporate these anatomical features may reveal insight into the interplay between top-down and bottom-up processing in the face network—a topic that has been elusive thus far.

In addition to macroscopic anatomical features such as white matter connectivity, we believe there is also utility in incorporating microanatomical features into computational models. For example, cytoarchitectonic differences between pFus- and mFus-faces suggest that these regions have different neural hardware that may be optimized to perform different types of computations. These data therefore suggest that computational models should not necessarily implement a single generic neural computation or filter type that is duplicated across processing stages. Instead, there are likely different specialized computations occurring at different stages of processing. Furthermore, an ambitious and interesting direction for future work is to

forge explicit links between anatomical properties such as cytoarchitecture, receptor architecture, myeloarchitecture, cell types, and microcircuit connections relative to the computational properties of neurons in the ventral face network.

1.4.2 What Can Deep Convolutional Networks Inform About Computational Strategies of the Brain?

Another promising avenue for future research will be to incorporate recent neuroscience findings into computational models implementing deep convolutional neural networks (CNNs; [41, 67, 85, 121, 136]), which have seen significant recent advancement. In brief, deep CNNs are neural networks composed from a series of stacked layers, in which the first layer performs operations on the input image and subsequent layers perform operations on the output of the prior layer. As such, CNNs have a feedforward architecture. Additionally, layers in a CNN typically alternate between layers performing linear operations (e.g., convolution), layers performing nonlinear operations (e.g., ReLU), and pooling layers. After several stacked layers of this sort, there are often one or more fully connected layers. The layers that perform linear operations typically contain multiple arrays of filters in which each filter performs an operation on a local region in the visual input (akin to computations by RFs in the human visual system), and the same filters are repeated across locations tiling the entire visual field. Additionally, filters in subsequent layers pool information from a local neighborhood from the prior layer, which yields an overall increase in pooling moving up the CNN hierarchy. Once the architecture of the deep CNN is built, the network is then trained to perform a task (e.g., face recognition). During training, the weights of the connections between layers are altered typically using a backpropagation algorithm [87, 127], which functions to reduce the error between the network output and the desired answer. After training, the weights are no longer changed and the processing in deep CNNs is strictly feedforward.

Deep CNN architectures are appealing because (1) they are inspired by architectural features of the ventral visual stream [41, 121, 136], (2) their performance reaches human-like performance in complex object recognition tasks [83, 177] including face recognition [145], and (3) they can predict—to a noteworthy level of accuracy—experimentally measured responses in the primate and human ventral stream [16, 33, 55, 67, 78, 86, 178]. We suggest that CNNs can be used as a tool (e.g., through simulations and analysis) to understand the computations being performed within the face network. However, as a first step, it is important to consider in what aspects the artificial architecture of deep CNNs is similar to or different than the neurobiological architecture of the human ventral visual stream.

Several aspects of deep CNNs are similar to the architecture of the human ventral visual stream (Table 1.1). For example, both systems implement: (1) computations by local filters, (2) hierarchical processing across a series of stages, (3) feedforward

Table 1.1 Computational hypotheses generated from comparisons between deep convolutional neural networks (CNNs) and neural architecture

Functional/ architectural property	Deep CNN	Human Brain	Hypothesized Computational Value
pRFs/filters			
Local computations	✔	✔	Parallel processing Distributed information
pRF/filter size increases along hierarchy	✔	✔	Useful features Invariance
pRF/filter size increases with eccentricity	✖	✔	Solution to limited brain size
Dynamic pRFs	✖	✔	Task-optimized processing
Representations			
Spatial topography	✖	✔	Efficiency and speed
Clustering of neurons with similar features	✖	✔	Reduce wiring length
Neural Hardware			
Differential neural across regions	✖?	✔?	Computation-optimized hardware
Connections			
Hierarchical connections	✔	✔	Hierarchical processing
Redundant connections	✔	✔	Resiliency to loss/damage
Bypass connections	✖	✔	Resiliency/speed
Top-down connections	during learning	✔	Learning/attentional gating

computations, (4) linear–nonlinear operations, (5) redundant connections, and (6) modification of weights during training.

While deep CNN architectures are broadly "neurally inspired" by neurobiological architectural features of the ventral visual stream, they also differ from the neural architecture in the human brain in several fundamental ways (Table 1.1). As examples, we note four differences between current deep CNNs such as AlexNet [85]) or FaceNet [131] and what is known about the ventral visual stream. First, after the training stage, filters in CNNs are fixed rather than dynamic, but in the brain, pRFs are dynamic and can be modified by top-down attention. Second, filters in CNNs are identical in size across the visual field, but in the brain, pRFs are larger in the periphery than the center of gaze. Third, in a given layer of a CNN there is no significance to the spatial arrangement of filters in a given layer, but the visual system

exhibits a spatial topography of representations across spatial scales of the cortical sheet that is reproducible across individuals. Fourth, in standard CNNs, beyond the training stage, there is typically no influence of top-down connections or bypass routes, but both types of connections are present and are used by the brain. By modifying the CNN architecture in order to implement features that more accurately reflect neurobiologically plausible brain architectures, it might be possible to explore and better understand the computational benefits of specific neurobiological features (Table 1.1 – right column). Below, we give examples of three such tests.

First, the computational architecture of most current deep CNNs such as AlexNet [85] or FaceNet [131] is strictly serial, which does not respect the biological reality that the ventral face network contains bypass routes that skip stages of the processing hierarchy. Therefore, incorporating recent empirical findings of bypass routes into the architectures of CNNs may (1) make CNNs more accurate models of the human visual system and (2) could advance our understanding regarding hypothesized benefits of specific architecture features such as bypass routes. For example, the hypothesis that bypass connections provide resiliency to cortical damage could be tested by comparing the effect of a virtual lesion to an intermediate layer of a strictly serial neural network versus a virtual lesion in a non-serial neural network containing bypass connections.

Second, deep CNNs contain fixed filters (beyond the training stage) and do not incorporate top-down connections from other processing stages. Consequently, processing in deep CNNs is purely stimulus-driven and therefore, does not account for known empirical effects of task and attention on responses in the ventral face network, such as attention-modulated pRFs. Thus, one could adapt a present deep CNN like FaceNet [131] to include top-down connections in order to test whether this addition explains task and attention effects on pRF properties and if so, if it improves the efficiency and/or performance of the network.

Third, as described in Sect. 1.2.1, there is a regular spatial topography of functional representations in the human brain. This topography is evident in all stages of the visual processing hierarchy—from an orderly structure of pRF arrangement across the cortex generating spatial maps of the visual field in early and intermediate visual areas [133, 161, 163], to the object form topography [22, 59] and regularity of face-selective regions relative to macroanatomical landmarks [165, 170] in ventral temporal cortex. In contrast to the spatial regularity of fine-scale functional regions and large-scale representations in the brain, there is (1) no spatial organization to nodes in a particular layer of a CNN and (2) neurobiological costs (such as wiring length) are not explicitly accounted for by CNNs, whereas biological systems are affected by those costs. Thus, future research can examine (1) what architectonical constraints need to be added to a deep CNN for it to develop a spatial structure (perhaps guided by the cytoarchitectonic and connectivity structure of face-selective regions summarized in Sect. 1.2 and (2) what computational benefit does cortical organization provide? For example, does spatial topography enhance CNN speed or efficiency?

1.5 Conclusions

As described in this chapter, neuroimaging research has advanced our understanding regarding the functional architecture of the human ventral face network. For example, the scale of our understanding of the spatial arrangement of regions in the ventral face network has improved from centimeters to millimeters. Recent research has linked this consistent spatial arrangement of the ventral face network to underlying microanatomical features such as cytoarchitecture, as well as white matter connectivity. Mechanistically, the development of new methods deriving population receptive fields has begun to elucidate computational principles of the ventral face network. While there are key questions that remain unanswered, these new research directions have opened exciting new opportunities for understanding how the functional neuroanatomical features of the face network contribute to computations underlying human face perception. Importantly, incorporating these recent findings in up-to-date computational models will further advance the field by providing enhanced understanding of the computational benefits of specific implementational features of the human brain by integrating such features into state-of-the-art deep convolutional neural network architectures.

Acknowledgements We thank Jesse Gomez for comments and edits of an earlier draft. This research has been funded by the following grants: NSF GRFP DGE-114747, 1RO1EY02231801-A1, and 1R01EY02391501-A1. Portions of this chapter were previously included within a paper published in the Annual Reviews of Vision Science, 2017 (http://www.annualreviews.org/doi/10.1146/annurev-vision-102016-061214).

Appendix: Abbreviations and Definitions

Collateral sulcus (CoS): a primary sulcus in human ventral temporal cortex; the medial boundary of the fusiform gyrus

Cytoarchitecture: cellular organization across the six-layered cortical ribbon; a property used to parcellate brain areas from one another

Developmental prosopagnosia (DP): an impairment in recognizing faces despite normal vision, intelligence, and socio-cognitive abilities and no history of brain damage

Fusiform face area (FFA): once considered a homogenous face-selective area, it contains (at least) two cytoarchitectonically and functionally distinct components

Fusiform gyrus (FG): a hominoid-specific macroanatomical structure in ventral temporal cortex that contains (at least) four cytoarchitectonic areas and multiple functional regions

FG1–4: Labels for four cytoarchitectonic areas in the fusiform gyrus and neighboring sulci

Inferior occipital gyrus (IOG): a gyrus that is posterior-lateral to the fusiform gyrus; considered the first processing stage of the ventral face network

Mid-fusiform sulcus (MFS): a shallow, longitudinal sulcus bisecting the fusiform gyrus; a landmark identifying cytoarchitectonic and functional boundaries

mFus-faces/FFA-2: a face-selective region overlapping the anterior-lateral tip of the mid-fusiform sulcus, located within cytoarchitectonic area FG4, and 1–1.5 cm anterior to pFus-faces/FFA-1

Occipito-temporal sulcus (OTS): a primary sulcus in human ventral temporal cortex; the lateral boundary of the fusiform gyrus

Parahippocampal place area (PPA): a place-selective region in the collateral sulcus and parahippocampal cortex

pFus-faces/FFA-1: a face-selective region typically overlapping the posterior-lateral tip of the mid-fusiform sulcus, located within cytoarchitectonic area FG2, and 1–1.5 cm posterior to mFus-faces/FFA-2

Population receptive field (pRF): in fMRI, the region of visual space that stimulates a voxel

Receptive field (RF): the region of the visual field which elicits a response from a neuron

Ventral Temporal Cortex (VTC): a cortical expanse spanning the inferior aspect of the temporal lobe containing high-level visual regions involved in "what" perception and recognition.

References

1. T. Allison, H. Ginter, G. McCarthy, A.C. Nobre, A. Puce et al., Face recognition in human extrastriate cortex. J. Neurophysiol. **71**, 821–825 (1994)
2. T. Allison, G. McCarthy, A. Nobre, A. Puce, A. Belger, Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. Cereb. Cortex **4**, 544–554 (1994)
3. T. Allison, A. Puce, D.D. Spencer, G. McCarthy, Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb. Cortex **9**, 415–430 (1999)
4. T.J. Andrews, M.P. Ewbank, Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage **23**, 905–913 (2004)
5. S. Anzellotti, S.L. Fairhall, A. Caramazza, Decoding representations of face identity that are tolerant to rotation. Cereb. Cortex **24**, 1988–1995 (2014)
6. G. Avidan, M. Behrmann, Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. Curr. Biol. **19**, 1146–1150 (2009)
7. G. Avidan, I. Levy, T. Hendler, E. Zohary, R. Malach, Spatial vs. object specific attention in high-order visual areas. Neuroimage **19**, 308–318 (2003)
8. G. Avidan, U. Hasson, R. Malach, M. Behrmann, Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. J. Cogn. Neurosci. **17**, 1150–1167 (2005)

9. V. Axelrod, G. Yovel, Hierarchical processing of face viewpoint in human visual cortex. J. Neurosci. **32**, 2442–2452 (2012)
10. V. Axelrod, G. Yovel, The challenge of localizing the anterior temporal face area: a possible solution. Neuroimage **2013**(81), 371–380 (2013)
11. M. Behrmann, G. Avidan, Congenital prosopagnosia: face-blind from birth. Trends Cogn. Sci. **9**, 180–187 (2005)
12. M. Behrmann, D.C. Plaut, Distributed circuits, not circumscribed centers, mediate visual recognition. Trends Cogn. Sci. **17**, 210–219 (2013)
13. M.G. Berman, J. Park, R. Gonzalez, T.A. Polk, A. Gehrke et al., Evaluating functional localizers: the case of the FFA. Neuroimage **50**, 56–71 (2010)
14. V. Bruce, A. Young, Understanding face recognition. Br. J. Psychol. **77**(Pt 3), 305–327 (1986)
15. L. Bugatus, K.S. Weiner, K. Grill-Spector, Task differentially modulates the spatial extent of category-selective regions across anatomical locations. Neuroimage (2017) (in press)
16. C.F. Cadieu, H. Hong, D.L. Yamins, N. Pinto, D. Ardila et al., Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLoS Comput. Biol. **10**, e1003963 (2014)
17. A.J. Calder, A.W. Young, Understanding the recognition of facial identity and facial expression. Nat. Rev. Neurosci. **6**, 641–651 (2005)
18. A.J. Calder, J.D. Beaver, J.S. Winston, R.J. Dolan, R. Jenkins et al., Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. Curr. Biol. **17**, 20–25 (2007)
19. T. Carlson, H. Hogendoorn, H. Fonteijn, F.A. Verstraten, Spatial coding and invariance in object-selective cortex. Cortex **47**, 14–22 (2011)
20. J. Caspers, K. Zilles, S.B. Eickhoff, A. Schleicher, H. Mohlberg, K. Amunts, Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. Brain Struct. Funct. **218**, 511–526 (2013)
21. J.A. Collins, I.R. Olson, Beyond the FFA: the role of the ventral anterior temporal lobes in face processing. Neuropsychologia **61**, 65–79 (2014)
22. A.C. Connolly, J.S. Guntupalli, J. Gors, M. Hanke, Y.O. Halchenko et al., The representation of biological classes in the human brain. J. Neurosci. **32**, 2608–2618 (2012)
23. T. Cukur, A.G. Huth, S. Nishimoto, J.L. Gallant, Functional subdomains within human FFA. J. Neurosci. **33**, 16748–16766 (2013)
24. N. Davidenko, D.A. Remus, K. Grill-Spector, Face-likeness and image variability drive responses in human face-selective ventral regions. Hum. Brain Mapp. **33**, 2234–2249 (2012)
25. I. Davidesco, M. Harel, M. Ramot, U. Kramer, S. Kipervasser et al., Spatial and object-based attention modulates broadband high-frequency responses across the human visual cortical hierarchy. J. Neurosci. **33**, 1228–1240 (2013)
26. I. Davidesco, E. Zion-Golumbic, S. Bickel, M. Harel, D.M. Groppe et al., Exemplar selectivity reflects perceptual similarities in the human fusiform cortex. Cereb. Cortex **24**, 1879–1893 (2014)
27. B. de Haas, D.S. Schwarzkopf, I. Alvarez, R.P. Lawson, L. Henriksson et al., Perception and processing of faces in the human brain is tuned to typical feature locations. J. Neurosci. **36**, 9289–9302 (2016)
28. B. Duchaine, K. Nakayama, The Cambridge face memory test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia **44**, 576–585 (2006)
29. B. Duchaine, G. Yovel, A revised neural framework for face processing. Annu. Rev. Vis. Sci **1**, 393–416 (2015)
30. B.C. Duchaine, K. Nakayama, Developmental prosopagnosia: a window to content-specific face processing. Curr. Opin. Neurobiol. **16**, 166–173 (2006)
31. S.O. Dumoulin, B.A. Wandell, Population receptive field estimates in human visual cortex. Neuroimage **39**, 647–660 (2008)
32. T. Egner, J.M. Monti, C. Summerfield, Expectation and surprise determine neural population responses in the ventral visual stream. J. Neurosci. **30**, 16601–16608 (2010)

33. M. Eickenberg, A. Gramfort, G. Varoquaux, B. Thirion, Seeing it all: convolutional network layers map the function of the human visual system. Neuroimage (2016)
34. M.P. Ewbank, T.J. Andrews, Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. Neuroimage **40**, 1857–1870 (2008)
35. F. Fang, S. He, Cortical responses to invisible objects in the human dorsal and ventral pathways. Nat. Neurosci. **8**, 1380–1385 (2005)
36. R. Farivar, O. Blanke, A. Chaudhuri, Dorsal-ventral integration in the recognition of motion-defined unfamiliar faces. J. Neurosci. **29**, 5336–5342 (2009)
37. D.J. Felleman, D.C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex **1**, 1–47 (1991)
38. B. Fischl, M.I. Sereno, R.B. Tootell, A.M. Dale, High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. **8**, 272–284 (1999)
39. W. Freiwald, B. Duchaine, G. Yovel, Face processing systems: from neurons to real-world social perception. Annu. Rev. Neurosci. **39**, 325–346 (2016)
40. M.A. Frost, R. Goebel, Measuring structural-functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. Neuroimage **59**, 1369–1381 (2012)
41. K. Fukushima, Neocognitron: a hierarchical neural network capable of visual pattern recognition. Neural Netw. **1**, 119–130 (1982)
42. I. Gauthier, P. Skudlarski, J.C. Gore, A.W. Anderson, Expertise for cars and birds recruits brain areas involved in face recognition. Nat. Neurosci. **3**, 191–197 (2000)
43. S. Gilaie-Dotan, R. Malach, Sub-exemplar shape tuning in human face-related areas. Cereb. Cortex **17**, 325–338 (2007)
44. S. Gilaie-Dotan, H. Gelbard-Sagiv, R. Malach, Perceptual shape sensitivity to upright and inverted faces is reflected in neuronal adaptation. Neuroimage **50**, 383–395 (2010)
45. J. Gomez, F. Pestilli, N. Witthoft, G. Golarai, A. Liberman et al., Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. Neuron **85**, 216–227 (2015)
46. C. Gratton, K.K. Sreenivasan, M.A. Silver, M. D'Esposito, Attention selectively modifies the representation of individual faces in the human brain. J. Neurosci. **33**, 6979–6989 (2013)
47. K. Grill-Spector, R. Malach, fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst) **107**, 293–321 (2001)
48. K. Grill-Spector, K.S. Weiner, The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. **15**, 536–548 (2014)
49. K. Grill-Spector, T. Kushnir, S. Edelman, G. Avidan, Y. Itzchak, R. Malach, Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron **24**, 187–203 (1999)
50. K. Grill-Spector, N. Knouf, N. Kanwisher, The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. **7**, 555–562 (2004)
51. K. Grill-Spector, R. Henson, A. Martin, Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. **10**, 14–23 (2006)
52. C.G. Gross, J. Sergent, Face recognition. Curr. Opin. Neurobiol. **2**, 156–161 (1992)
53. C.G. Gross, D.B. Bender, C.E. Rocha-Miranda, Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science **166**, 1303–1306 (1969)
54. M. Gschwind, G. Pourtois, S. Schwartz, D. Van De Ville, P. Vuilleumier, White-matter connectivity between face-responsive regions in the human brain. Cereb. Cortex **22**, 1564–1576 (2012)
55. U. Guclu, M.A. van Gerven, Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. J. Neurosci. **35**, 10005–10014 (2015)
56. R.J. Harris, G.E. Rice, A.W. Young, T.J. Andrews, Distinct but overlapping patterns of response to words and faces in the fusiform gyrus. Cereb. Cortex **26**, 3161–3168 (2016)
57. U. Hasson, I. Levy, M. Behrmann, T. Hendler, R. Malach, Eccentricity bias as an organizing principle for human high-order object areas. Neuron **34**, 479–490 (2002)

58. J.V. Haxby, E.A. Hoffman, M.I. Gobbini, The distributed human neural system for face perception. Trends Cogn. Sci. **4**, 223–233 (2000)
59. J.V. Haxby, M.I. Gobbini, M.L. Furey, A. Ishai, J.L. Schouten, P. Pietrini, Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science **293**, 2425–2430 (2001)
60. C.C. Hemond, N.G. Kanwisher, H.P. Op de Beeck, A preference for contralateral stimuli in human object- and face-selective cortex. PLoS One **2**, e574 (2007)
61. L. Henriksson, M. Mur, N. Kriegeskorte, Faciotopy-A face-feature map with face-like topology in the human occipital face area. Cortex **72**, 156–167 (2015)
62. G. Holmes, Disturbances of vision by cerebral lesions. Br. J. Ophthalmol. **2**, 353–384 (1918)
63. D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. **160**, 106–154 (1962)
64. J.B. Hutchinson, M.R. Uncapher, K.S. Weiner, D.W. Bressler, M.A. Silver et al., Functional heterogeneity in posterior parietal cortex across attention and episodic memory retrieval. Cereb. Cortex **24**, 49–66 (2012)
65. A. Ishai, L.G. Ungerleider, A. Martin, J.V. Haxby, The representation of objects in the human occipital and temporal cortex. J. Cogn. Neurosci. **12**(Suppl 2), 35–51 (2000)
66. C. Jacques, N. Witthoft, K.S. Weiner, B.L. Foster, V. Rangarajan et al., Corresponding ECoG and fMRI category-selective signals in human ventral temporal cortex. Neuropsychologia **83**, 14–28 (2016)
67. X. Jiang, E. Rosen, T. Zeffiro, J. Vanmeter, V. Blanz, M. Riesenhuber, Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. Neuron **50**, 159–172 (2006)
68. J. Jonas, S. Frismand, J.P. Vignal, S. Colnat-Coulbois, L. Koessler et al., Right hemispheric dominance of visual phenomena evoked by intracerebral stimulation of the human visual cortex. Hum. Brain Mapp. **35**, 3360–3371 (2014)
69. J. Jonas, B. Rossion, J. Krieg, L. Koessler, S. Colnat-Coulbois et al., Intracerebral electrical stimulation of a face-selective area in the right inferior occipital cortex impairs individual face discrimination. Neuroimage **99**, 487–497 (2014)
70. J. Jonas, C. Jacques, J. Liu-Shuang, H. Brissart, S. Colnat-Coulbois et al., A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. Proc. Natl. Acad. Sci. USA **113**, E4088–E4097 (2016)
71. N. Kanwisher, Domain specificity in face perception. Nat. Neurosci. **3**, 759–763 (2000)
72. N. Kanwisher, J. McDermott, M.M. Chun, The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. **17**, 4302–4311 (1997)
73. N. Kanwisher, F. Tong, K. Nakayama, The effect of face inversion on the human fusiform face area. Cognition **68**, B1–B11 (1998)
74. K.N. Kay, J.D. Yeatman, Bottom-up and top-down computations in high-level visual cortex. Elife. 2017 Feb 22;6. pii: e22341 (2017)
75. K.N. Kay, T. Naselaris, R.J. Prenger, J.L. Gallant, Identifying natural images from human brain activity. Nature **452**, 352–355 (2008)
76. K.N. Kay, J. Winawer, A. Mezer, B.A. Wandell, Compressive spatial summation in human visual cortex. J. Neurophysiol. **110**, 481–494 (2013)
77. K.N. Kay, K.S. Weiner, K. Grill-Spector, Attention reduces spatial uncertainty in human ventral temporal cortex. Curr. Biol. **25**, 595–600 (2015)
78. S.M. Khaligh-Razavi, N. Kriegeskorte, Deep supervised, but not unsupervised, models may explain IT cortical representation. PLoS Comput. Biol. **10**, e1003915 (2014)
79. T.C. Kietzmann, J.D. Swisher, P. Konig, F. Tong, Prevalence of selectivity for mirror-symmetric views of faces in the ventral and dorsal visual pathways. J. Neurosci. **32**, 11763–11772 (2012)
80. M. Kim, M. Ducros, T. Carlson, I. Ronen, S. He et al., Anatomical correlates of the functional organization in the human occipitotemporal cortex. Magn. Reson. Imaging **24**, 583–590 (2006)

81. B.P. Klein, B.M. Harvey, S.O. Dumoulin, Attraction of position preference by spatial attention throughout human visual cortex. Neuron **84**, 227–237 (2014)
82. J. Konorski, *Integrative Activity of the Brain. An Interdisciplinary Approach* (The University of Chicago Press, Chicago, 1967)
83. N. Kriegeskorte, Deep neural networks: a new framework for modeling biological vision and brain information processing. Annu. Rev. Vis. Sci. **1**, 417–446 (2015)
84. N. Kriegeskorte, E. Formisano, B. Sorger, R. Goebel, Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. USA **104**, 20600–20605 (2007)
85. A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks. Presented at Neural Information Processing Systems (NIPS) (2012)
86. J. Kubilius, S. Bracci, H.P. Op de Beeck, Deep neural networks as a computational model for human shape sensitivity. PLoS Comput. Biol. **12**, e1004896 (2016)
87. Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard et al., Neural Information Processing (1989)
88. Y. Lee, B. Duchaine, H.R. Wilson, K. Nakayama, Three cases of developmental prosopagnosia from one family: detailed neuropsychological and psychophysical investigation of face processing. Cortex **46**, 949–964 (2010)
89. I. Levy, U. Hasson, G. Avidan, T. Hendler, R. Malach, Center-periphery organization of human object areas. Nat. Neurosci. **4**, 533–539 (2001)
90. G. Loffler, G. Yourganov, F. Wilkinson, H.R. Wilson, fMRI evidence for the neural representation of faces. Nat. Neurosci. **8**, 1386–1390 (2005)
91. G.R. Loftus, E.M. Harley, Why is it easier to identify someone close than far away? Psychon. Bull. Rev. **12**, 43–65 (2005)
92. S. Lorenz, K.S. Weiner, J. Caspers, H. Mohlberg, A. Schleicher et al., Two new cytoarchitectonic areas on the human mid-fusiform gyrus. Cereb. Cortex (2015)
93. A. Martin, C.L. Wiggs, L.G. Ungerleider, J.V. Haxby, Neural correlates of category-specific knowledge. Nature **379**, 649–652 (1996)
94. G. McCarthy, A. Puce, J.C. Gore, T. Allison, Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. **9**, 605–610 (1997)
95. G. McCarthy, A. Puce, A. Belger, T. Allison, Electrophysiological studies of human face perception. II: response properties of face-specific potentials generated in occipitotemporal cortex. Cereb. Cortex **9**, 431–444 (1999)
96. E. McKone, Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. Vis. Res. **49**, 268–283 (2009)
97. K. Moutoussis, S. Zeki, The relationship between cortical activation and perception investigated with invisible stimuli. Proc. Natl. Acad. Sci. USA **99**, 9527–9532 (2002)
98. M. Mur, D.A. Ruff, J. Bodurka, P. De Weerd, P.A. Bandettini, N. Kriegeskorte, Categorical, yet graded-single-image activation profiles of human category-selective cortical regions. J. Neurosci. **32**, 8649–8662 (2012)
99. S. Nasr, N. Liu, K.J. Devaney, X. Yue, R. Rajimehr et al., Scene-selective cortical regions in human and nonhuman primates. J. Neurosci. **31**, 13771–13785 (2011)
100. V. Natu, M.A. Barnett, T. Hartley, J. Gomez, A. Stigliani, K. Grill-Spector, Development of neural sensitivity to face identity correlates with perceptual discriminability. J. Neurosci. **36**, 10893–10907 (2016)
101. V.S. Natu, A.J. O'Toole, Spatiotemporal changes in neural response patterns to faces varying in visual familiarity. Neuroimage **108**, 151–159 (2015)
102. V.S. Natu, F. Jiang, A. Narvekar, S. Keshvari, V. Blanz, A.J. O'Toole, Dissociable neural patterns of facial identity across changes in viewpoint. J. Cogn. Neurosci. **22**, 1570–1582 (2010)
103. A. Nestor, D.C. Plaut, M. Behrmann, Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proc. Natl. Acad. Sci. USA **108**, 9998–10003 (2011)
104. K.M. O'Craven, P.E. Downing, N. Kanwisher, fMRI evidence for objects as the units of attentional selection. Nature **401**, 584–587 (1999)

105. J. Parvizi, C. Jacques, B.L. Foster, N. Withoft, V. Rangarajan et al., Electrical stimulation of human fusiform face-selective regions distorts face perception. J. Neurosci. **32**, 14915–14920 (2012)
106. M.V. Peelen, P.E. Downing, Within-subject reproducibility of category-specific visual activation with functional MRI. Hum. Brain Mapp. **25**, 402–408 (2005)
107. K.A. Pelphrey, N.J. Sasson, J.S. Reznick, G. Paul, B.D. Goldman, J. Piven, Visual scanning of faces in autism. J Autism Dev. Disord. **32**, 249–261 (2002)
108. M.A. Pinsk, M. Arcaro, K.S. Weiner, J.F. Kalkus, S.J. Inati et al., Neural representations of faces and body parts in macaque and human cortex: a comparative FMRI study. J. Neurophysiol. **101**, 2581–2600 (2009)
109. D. Pitcher, V. Walsh, G. Yovel, B. Duchaine, TMS evidence for the involvement of the right occipital face area in early face processing. Curr. Biol. **17**, 1568–1573 (2007)
110. D. Pitcher, D.D. Dilks, R.R. Saxe, C. Triantafyllou, N. Kanwisher, Differential selectivity for dynamic versus static information in face-selective cortical regions. Neuroimage **56**, 2356–2363 (2011)
111. D. Pitcher, V. Walsh, B. Duchaine, The role of the occipital face area in the cortical face perception network. Exp. Brain Res. **209**, 481–493 (2011)
112. D. Pitcher, T. Goldhaber, B. Duchaine, V. Walsh, N. Kanwisher, Two critical and functionally distinct stages of face and body perception. J. Neurosci. **32**, 15877–15885 (2012)
113. E. Privman, Y. Nir, U. Kramer, S. Kipervasser, F. Andelman et al., Enhanced category tuning revealed by intracranial electroencephalograms in high-order human visual areas. J. Neurosci. **27**, 6234–6242 (2007)
114. A. Puce, T. Allison, J.C. Gore, G. McCarthy, Face-sensitive regions in human extrastriate cortex studied by functional MRI. J. Neurophysiol. **74**, 1192–1199 (1995)
115. A. Puce, T. Allison, S. Bentin, J.C. Gore, G. McCarthy, Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci. **18**, 2188–2199 (1998)
116. A. Puce, T. Allison, G. McCarthy, Electrophysiological studies of human face perception. III: effects of top-down processing on face-specific potentials. Cereb. Cortex **9**, 445–458 (1999)
117. J.A. Pyles, T.D. Verstynen, W. Schneider, M.J. Tarr, Explicating the face perception network with white matter connectivity. PLoS One **8**, e61611 (2013)
118. R. Rajimehr, J.C. Young, R.B. Tootell, An anterior temporal face patch in human cortex, predicted by macaque maps. Proc. Natl. Acad. Sci. USA **106**, 1995–2000 (2009)
119. V. Rangarajan, D. Hermes, B.L. Foster, K.S. Weiner, C. Jacques et al., Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. J. Neurosci. **34**, 12828–12836 (2014)
120. J.H. Reynolds, D.J. Heeger, The normalization model of attention. Neuron **61**, 168–185 (2009)
121. M. Riesenhuber, T. Poggio, Hierarchical models of object recognition in cortex. Nat. Neurosci. **2**, 1019–1025 (1999)
122. M. Rosenke, K.S. Weiner, M.A. Barnett, K. Zilles, K. Amunts, R. Goebel, K. Grill-Spector, A cross-validated cytoarchitectonic atlas of the human ventral visual stream. NeuroImage **pii: S1053–8119**(17), 30151–30159 (2017)
123. B. Rossion, Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. Neuroimage **40**, 423–426 (2008)
124. B. Rossion, A. Boremanse, Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. J. Vis. **11**, 1–18 (2011)
125. B. Rossion, R. Caldara, M. Seghier, A.M. Schuller, F. Lazeyras, E. Mayer, A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain **126**, 2381–2395 (2003)
126. P. Rotshtein, R.N. Henson, A. Treves, J. Driver, R.J. Dolan, Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. Nat. Neurosci. **8**, 107–113 (2005)
127. D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, *Cognitive Modeling* (1988)

128. Z.M. Saygin, D.E. Osher, K. Koldewyn, G. Reynolds, J.D. Gabrieli, R.R. Saxe, Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. Nat. Neurosci. **15**, 321–327 (2012)
129. C. Schiltz, B. Sorger, R. Caldara, F. Ahmed, E. Mayer et al., Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. Cereb. Cortex **16**, 574–586 (2006)
130. C. Schiltz, L. Dricot, R. Goebel, B. Rossion, Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. J. Vis. **10**(25), 1–16 (2010)
131. F. Schroff, D. Kalenichenko, J. Philbin, FaceNet: a unified embedding for face recognition and clustering. in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR, 2015)* (2015), pp. 815–823
132. R.F. Schwarzlose, J.D. Swisher, S. Dang, N. Kanwisher, The distribution of category and location information across object-selective regions in human visual cortex. Proc. Natl. Acad. Sci. USA **105**, 4447–4452 (2008)
133. M.I. Sereno, A.M. Dale, J.B. Reppas, K.K. Kwong, J.W. Belliveau et al., Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science **268**, 889–893 (1995)
134. J. Sergent, J.L. Signoret, Functional and anatomical decomposition of face processing: evidence from prosopagnosia and PET study of normal subjects. Philos. Trans. R. Soc. Lond. B Biol. Sci. **335**, 55–61; discussion 61-2
135. J. Sergent, S. Ohta, B. MacDonald, Functional neuroanatomy of face and object processing. A positron emission tomography study. Brain **115**(Pt 1), 15–36 (1992)
136. T. Serre, A. Oliva, T. Poggio, A feedforward architecture accounts for rapid categorization. Proc. Natl. Acad. Sci. USA **104**, 6424–6429 (2007)
137. M.A. Silver, S. Kastner, Topographic maps in human frontal and parietal cortex. Trends Cogn. Sci. **13**, 488–495 (2009)
138. M.A. Silver, D. Ress, D.J. Heeger, Topographic maps of visual spatial attention in human parietal cortex. J. Neurophysiol. **94**, 1358–1371 (2005)
139. H.P. Snippe, J.J. Koenderink, Information in channel-coded systems: correlated receivers. Biol. Cybern. **67**, 183–190 (1992)
140. B. Sorger, R. Goebel, C. Schiltz, B. Rossion, Understanding the functional neuroanatomy of acquired prosopagnosia. Neuroimage **35**, 836–852 (2007)
141. T.C. Sprague, J.T. Serences, Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. Nat. Neurosci. **16**, 1879–1887 (2013)
142. J.K. Steeves, J.C. Culham, B.C. Duchaine, C.C. Pratesi, K.F. Valyear et al., The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. Neuropsychologia **44**, 594–609 (2006)
143. C. Summerfield, E.H. Trittschuh, J.M. Monti, M.M. Mesulam, T. Egner, Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. (2008)
144. J.D. Swisher, M.A. Halko, L.B. Merabet, S.A. McMains, D.C. Somers, Visual topography of human intraparietal sulcus. J. Neurosci. **27**, 5326–5337 (2007)
145. Y. Taigman, M. Yang, M. Ranzato, L. Wolf, in *The IEEE Conference on Computer Vision and Pattern Recognition* (CVPR, 2014), pp. 1701–1708
146. H. Takemura, A. Rokem, J. Winawer, J.D. Yeatman, B.A. Wandell, F. Pestilli, A major human white matter pathway between dorsal and ventral visual cortex. Cereb. Cortex **26**, 2205–2214 (2016)
147. T. Tallinen, J.Y. Chung, J.S. Biggins, L. Mahadevan, Gyrification from constrained cortical expansion. Proc. Natl. Acad. Sci. USA **111**, 12667–12672 (2014)
148. J.W. Tanaka, M.J. Farah, Parts and wholes in face recognition. Q. J. Exp. Psychol. A Hum. Exp. Psychol. **46**, 225–245 (1993)
149. I. Tavor, M. Yablonski, A. Mezer, S. Rom, Y. Assaf, G. Yovel, Separate parts of occipito-temporal white matter fibers are associated with recognition of faces and places. Neuroimage (2013)

150. F. Tong, K. Nakayama, J.T. Vaughan, N. Kanwisher, Binocular rivalry and visual awareness in human extrastriate cortex. Neuron **21**, 753–759 (1998)
151. F. Tong, K. Nakayama, M. Moscovitch, O. Weinrib, N. Kanwisher, Response properties of the human fusiform face area. Cogn. Neuropsychol. **17**, 257–280 (2000)
152. D. Tsao, M. Livingstone, Mechanisms of face perception. Annu. Rev. Neurosci. **31**, 411–437 (2008)
153. D.Y. Tsao, S. Moeller, W.A. Freiwald, Comparing face patch systems in macaques and humans. Proc. Natl. Acad. Sci. USA **105**, 19514–19519 (2008)
154. T. Valentine, Face-space models of face recognition, in *Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges*, ed. by M.J. Wenger, J.T. Townsend (Lawrence Erlbaum Associates Inc, Hillsdale, 2001)
155. G. Van Belle, P. De Graef, K. Verfaillie, T. Busigny, B. Rossion, Whole not hole: expert face recognition requires holistic perception. Neuropsychologia **48**, 2620–2629 (2010)
156. G. Van Belle, T. Busigny, P. Lefevre, S. Joubert, O. Felician et al., Impairment of holistic face perception following right occipito-temporal damage in prosopagnosia: converging evidence from gaze-contingency. Neuropsychologia **49**, 3145–3150 (2011)
157. D.C. Van Essen, J.L. Gallant, Neural mechanisms of form and motion processing in the primate visual system. Neuron **13**, 1–10 (1994)
158. D.C. Van Essen, C.H. Anderson, D.J. Felleman, Information processing in the primate visual system: an integrated systems perspective. Science **255**, 419–423 (1992)
159. P. Vuilleumier, R.N. Henson, J. Driver, R.J. Dolan, Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat. Neurosci. **5**, 491–499 (2002)
160. P. Vuilleumier, J.L. Armony, J. Driver, R.J. Dolan, Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nat. Neurosci. **6**, 624–631 (2003)
161. B.A. Wandell, J. Winawer, Imaging retinotopic maps in the human brain. Vis. Res. **51**, 718–737 (2011)
162. B.A. Wandell, J. Winawer, Computational neuroimaging and population receptive fields. Trends Cogn. Sci. **19**, 349–357 (2015)
163. B.A. Wandell, S.O. Dumoulin, A.A. Brewer, Visual field maps in human cortex. Neuron **56**, 366–383 (2007)
164. K. Weibert, T.J. Andrews, Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. Neuropsychologia **75**, 588–596 (2015)
165. K.S. Weiner, K. Grill-Spector, Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. Neuroimage **52**, 1559–1573 (2010)
166. K.S. Weiner, K. Grill-Spector, The improbable simplicity of the fusiform face area. Trends Cogn. Sci. **16**(5), 251–4 (2012)
167. K.S. Weiner, K. Grill-Spector, The evolution of face processing networks. Trends Cogn. Sci. **19**, 240–241 (2015)
168. K.S. Weiner, K. Zilles, The anatomical and functional specialization of the fusiform gyrus. Neuropsychologia **83**, 48–62 (2016)
169. K.S. Weiner, R. Sayres, J. Vinberg, K. Grill-Spector, fMRI-adaptation and category selectivity in human ventral temporal cortex: regional differences across time scales. J. Neurophysiol. **103**, 3349–3365 (2010)
170. K.S. Weiner, G. Golarai, J. Caspers, M.R. Chuapoco, H. Mohlberg et al., The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. Neuroimage **84**, 453–465 (2014)
171. K.S. Weiner, J. Jonas, J. Gomez, L. Maillard, H. Brissart et al., The face-processing network is resilient to focal resection of human visual cortex. J. Neurosci. **36**, 8425–8440 (2016)
172. K.S. Weiner, J.D. Yeatman, B.A. Wandell, The posterior arcuate fasciculus and the vertical occipital fasciculus. Cortex. 31 Mar 2016. pii: (16)30050-8 (2016). doi:10.1016/S0010-9452
173. K.S. Weiner, B.A. Barnett, S. Lorenz, J. Caspers, A. Stigliani et al., The cytoarchitecture of domain-specific regions in human high-level visual cortex. Cereb. Cortex (2017)
174. Y. Weiss, S. Edelman, M. Fahle, Models of perceptual learning in vernier hyperacuity. Neural Comput. **5**, 695–718 (1993)

175. J.S. Winston, R.N. Henson, M.R. Fine-Goulden, R.J. Dolan, fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. J. Neurophysiol. **92**, 1830–1839 (2004)
176. N. Witthoft, S. Poltoratski, M. Nguyen, G. Golarai, A. Liberman et al., Developmental prosopagnosia is associated with reduced spatial integration in the ventral visual cortex, in *bioRxiv* (2016)
177. D.L. Yamins, J.J. DiCarlo, Using goal-driven deep learning models to understand sensory cortex. Nat. Neurosci. **19**, 356–365 (2016)
178. D.L. Yamins, H. Hong, C.F. Cadieu, E.A. Solomon, D. Seibert, J.J. DiCarlo, Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. USA **111**, 8619–8624 (2014)
179. J.D. Yeatman, K.S. Weiner, F. Pestilli, A. Rokem, A. Mezer, B.A. Wandell, The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. Proc. Natl. Acad. Sci. USA **111**, E5214–E5223 (2014)
180. D.J. Yi, T.A. Kelley, R. Marois, M.M. Chun, Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. Brain Res. **1080**, 53–62 (2006)
181. G. Yovel, N. Kanwisher, Face perception: domain specific, not process specific. Neuron **44**, 889–898 (2004)
182. G. Yovel, N. Kanwisher, The neural basis of the behavioral face-inversion effect. Curr. Biol. **15**, 2256–2262 (2005)
183. X. Yue, B.S. Cassidy, K.J. Devaney, D.J. Holt, R.B. Tootell, Lower-level stimulus features strongly influence responses in the fusiform face area. Cereb. Cortex **21**, 35–47 (2011)
184. S. Zeki, S. Shipp, The functional logic of cortical connections. Nature **335**, 311–317 (1988)