



HAL
open science

Brief Overview of Functional Imaging Principles

Christophe N Habas, Giovanni de Marco

► **To cite this version:**

Christophe N Habas, Giovanni de Marco. Brief Overview of Functional Imaging Principles. The Neuroimaging of Brain Diseases, Springer; Springer International Publishing, pp.27-58, 2018, 10.1007/978-3-319-78926-2_2. hal-02310399

HAL Id: hal-02310399

<https://hal.parisnanterre.fr/hal-02310399>

Submitted on 10 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Brief Overview of Functional Imaging Principles

C. Habas ¹✉

Email chabas@15-20.fr

G. de Marco ²

¹ Service de NeuroImagerie, CHNO des 15-20, Paris, France

² Laboratoire ceRSM (EA-2931), Equipe Analyse du Mouvement Biomécanique, Physiologie et Imagerie, Université Paris Ouest, Nanterre, France

Abstract

The abstract is published online only. If you did not include a short abstract for the online version when you submitted the manuscript, the first paragraph or the first 10 lines of the chapter will be displayed here. If possible, please provide us with an informative abstract.

Functional imaging enables to detect and to localize brain areas specifically involved in networks subserving a given mental activity. The two main techniques used in routine consist in functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). We will focus on fMRI. fMRI is based on the local and transient increase of blood oxygenation in the cortical and deep nuclear microvasculature, caused by neuronal activation. This hemodynamic phenomenon called blood-oxygenation level-dependent (BOLD) response only indirectly reflects the amount of neuronal activity (Raichle ME, Proc Natl Acad Sci USA 93:765–772, 1998). This local blood hyperoxygenation produces microscopic magnetic field alterations measurable by appropriate T2*-weighted MRI sequence. A complex image and statistical post-processing is then applied to raw data to generate final “brain activation maps.” These functional maps can be acquired during active stimulation protocols, or “at rest” (resting-state functional connectivity, rsfMRI). fMRI and rsfMRI resort to different algorithmic processing: linear model-related algorithms and nonlinear, model-free, data-driven algorithms, respectively. This functional connectivity can be complemented by effective connectivity which seeks causality relationships between brain areas participating to a same circuit. Finally, advanced algorithms can also explore the topology of the neural networks and provide new and richer graph-theoretic related information, and machine learning will improve normal and aberrant structural and functional brain pattern recognition and classification.

Keywords

Functional MRI

Neurovascular coupling

BOLD

ASL

GLM

ICA

Brain resting state

Functional connectivity

Topology

Graph theory

Physiological and Physical Bases of BOLD Response

BOLD

Blood-oxygenation level-dependent (BOLD)

signal roots in local and transient hyperoxygenation exceeding neuronal metabolic demand of activated neurons and is due to vasodilation partly and feed-forwardly controlled by vasoactive molecules in relation to (inter-)neuronal activity and to energetic metabolism [1] and modulated by more distant neural influences.

Neurovascular Coupling with Glutamatergic Neurons At rest, the main neuronal consumption of energy produced by aerobic glycolysis (tricarboxylic acid cycle) is linked to the maintenance of gradients of ionic concentration around the plasmic membrane and intracellular biosynthesis. During synaptic activation, neuronal metabolism and, consequently, neuronal consumption/production of energy are increased mainly in relation with the recruitment of Na⁺/K⁺ ATPase pumps involved in gradient ion restoration. A tight neuron-glia-vascular coupling allows to provide sufficient amount of oxygen and glucose to neurons (and astrocytes) during sustained activity and during restoration of their energetic reserves and to get rid of metabolic potentially noxious by-products [2]. This coupling induces (1) a strong and nonlinear augmentation of cerebral blood flow (CBF) and (2) a moderate augmentation of cerebral blood volume (CBV), while a weak cerebral metabolic rate of oxygen utilization (CMRO₂) is observed [3]. This hemodynamic response is based on a vasodilation of arterioles and capillaries causing increased CBF and, in downstream veins, an increased speed of blood flow [2]. These vascular processes would rely on Ca²⁺-dependent release of vasodilators by neurons (or interneurons) and by the associated astrocytes whose metabotropic receptors are activated by the glutamate released in the synaptic cleft and whose end feet surround the microvasculature. These molecules can act on pericytes and smooth muscle cells. Several factors can contribute to the local vasodilation [4]: astrocytic extrusion of potassium, prostaglandins, and epoxyeicosatrienoic acids; neuronal liberation of vasoactive mediators such as nitric oxide (NO), prostaglandins, vasoactive intestinal peptide (VIP), adenosine, or adenine; and variation of blood pCO₂ and pH and lactate/pyruvate ratio. It is noteworthy that this vasodilation can be potentially counterbalanced by vasoconstrictors such as norepinephrine. Two other mechanisms can also participate to this functional hyperemia. First, some vasoactive substances may be released from distant subcortical afferents such as from the cholinergic substantia innominata, serotonergic raphe, dopaminergic ventral tegmental area, and noradrenergic locus coeruleus [2, 4]. Second, (retro-)propagation of vasodilation can occur in relation to endothelial direct action on smooth muscle cells via hyperpolarization-dependent [5] or calcium wave-dependent vasoactive factors [6]. This mechanism would explain why vasodilation can also spread to the arterioles toward the cortical surface [2].

Neurovascular Coupling with GABAergic Interneurons Stimulation of GABA-A receptors produces arteriolar dilation in the neocortex and hippocampus so that GABAergic interneurons may contribute to positive BOLD response [7]. However, several experiments have demonstrated that inhibitory GABAergic synapse could also induce arteriolar vasoconstriction and subsequent decrease in CBV and blood oxygenation, likely through neuropeptide corelease [4]. This hemodynamic response would partly explain negative BOLD response. However, interneurons, such as cerebellar basket and stellate cells, can also secrete molecules such as NO producing vasodilation, while GABA has no vasoactive effects in the cerebellum [8]. Finally, local neurovascular response represents summed and complex effects of both excitatory and inhibitory synapses.

Neuronal Metabolism of Glutamatergic Neurons Moreover, part of the glutamate delivered in the synaptic cleft is captured by contiguous astrocytes and transformed into glutamine before being released back to presynaptic neuron where it will be recycled [9, 10]. The glial sodium-glutamate cotransport stimulates anaerobic glycolysis which produces ATP necessary to extrude sodium through a sodium-potassium ATPase and lactate delivered to the neuron using monocarboxylate transporters. Lactate is then transformed into pyruvate by neuronal lactate dehydrogenase before undergoing oxidative phosphorylation, according to the lactate shuttle hypothesis. This mechanism would explain the observation that only a small fraction of blood oxygen is extracted, while a substantial uptake of glucose occurs.

Local Magnetic Field Owing to functional hyperemia, the surplus of diamagnetic oxyhemoglobin (mainly due to the increased CBV) flows in the local capillaries and venous system which become more diamagnetic [11]. Let's remind that, at rest, the difference of magnetic susceptibility between the paramagnetic deoxyhemoglobin-rich vascular compartment and the more diamagnetic surrounding tissue creates a microscopic field gradient around vessels causing variation of the precessional frequency with space and time. Moreover, randomly diffusing spins (hydrogen atoms of water molecules) due to thermal energy experience different amplitudes of the local magnetic field and exhibit subsequent accrual phase. However, this effect dominates mainly around the small vessels (capillaries and venules) as the diameter of the spin diffusion sphere on the order of 10 μm and the spatial spread of the gradient are approximatively of the same range. In larger vessels (arterioles and large venules), spins can be regarded as static in the perivascular space. It is noteworthy that dephasing also takes place in the intravascular compartment where diffusing spins undergo direct magnetic influence of deoxyhemoglobin because of their spatial closeness. Conversely, after neuronal activation, the increased concentration of oxyhemoglobin in vessels leads to decreased local field distortions and consequently lesser phase dispersion of water molecules. In summary, augmentation of CBF due to neuron-glia-vascular coupling is accompanied by a net increased oxygenation of the microvasculature, while CBV and oxygen consuming induce a counterbalancing but weaker augmentation of deoxyhemoglobin, so that the local magnetic field is significantly diminished. This change of local tissue oxygenation-related magnetism is recorded by fMRI as BOLD response.

BOLD Signal and fMRI After radio-frequency pulse, fMRI detects transversal magnetization whose decay depends upon static and fluctuating field inhomogeneities, including the deoxyhemoglobin-related field gradient, and on spin motion [11, 12]. These magnetic inhomogeneities contribute to faster loss of spin-phase coherence, and this effect is reflected in the $T2^*$ relaxation time which appears in the equation describing the MRI signal evolution in function of time $S(t)$ with a baseline signal equal to S_0 :

$$S(t) = S_0 \cdot \exp[-t/T2^*]$$

After neuronal activation, $T2^*$ increases to $T2^{*'}$ because of the decrease of intravascular and perivascular paramagnetic distortions of the main field exerted by deoxyhemoglobin transiently replaced by oxyhemoglobin. Therefore, the measured fMRI signal appears stronger. It can be easily calculated that the percentage of change of the signal between the rest baseline and the activated state, which represents the BOLD signal, is approximatively proportional to $[1/T2^* - 1/T2^{*'}] TE$, where TE refers to echo time of the MRI sequence. Several biophysical models have been developed to express the relaxation time $T2^*$ in terms of local field and vascular parameters. For instance, it has been shown that [12, 13]:

$$1/T2^* \approx k \cdot ([dHb] \cdot B_0)^\beta V$$

where k designates a constant, $[dHb]$ the concentration of deoxyhemoglobin which determines the perivascular difference of magnetic susceptibility, B_0 the main magnetic field strength, V the blood volume fraction, and $\beta = 1.5$. More complex and accurate models integrate the dependency of $[dHb]$ on CBF, CBV, and the rate of oxygen extraction. For instance, one possible model for the BOLD response ΔS can be approached by the following nonlinear (in CBF) equation [14]:

$$\Delta S(t) = A \cdot (1 - 1/n - \alpha_V) \cdot (1 - CBF(t) / CBF_{baseline})$$

where the complex factor A is mainly proportional to $[dHb]$ (and consequently to hematocrit) and O_2 extraction during the baseline state, n is the fractional changes in CBF, and α_V refers to the venous volume changes. It is important to emphasize that the BOLD signal in each pixel is strongly influenced by baseline values of CBF and CBV, for instance. Therefore, BOLD signal does not only depend on neuronal activation but also on regional-specific microvascular anatomy and physiological changes related to age, medication, or concomitant pathology.

As mentioned above, negative BOLD response can also be observed. Kim and Ogawa [15] summarized several possible explanations of stimulus-induced negative BOLD signals: decreased CBF due to neuronal inhibition, or regional flow reallocation, and increased O_2 consumption without augmented CBF. Furthermore, they reported Shih et al.'s study [16] showing subcortical negative BOLD response (in rats) mediated by focal release of vasoconstrictors decreasing CBV despite nociceptive-induced neuronal activation. Therefore, negative BOLD signal must be carefully interpreted and must not be regarded as exclusive signature of neuronal inhibition.

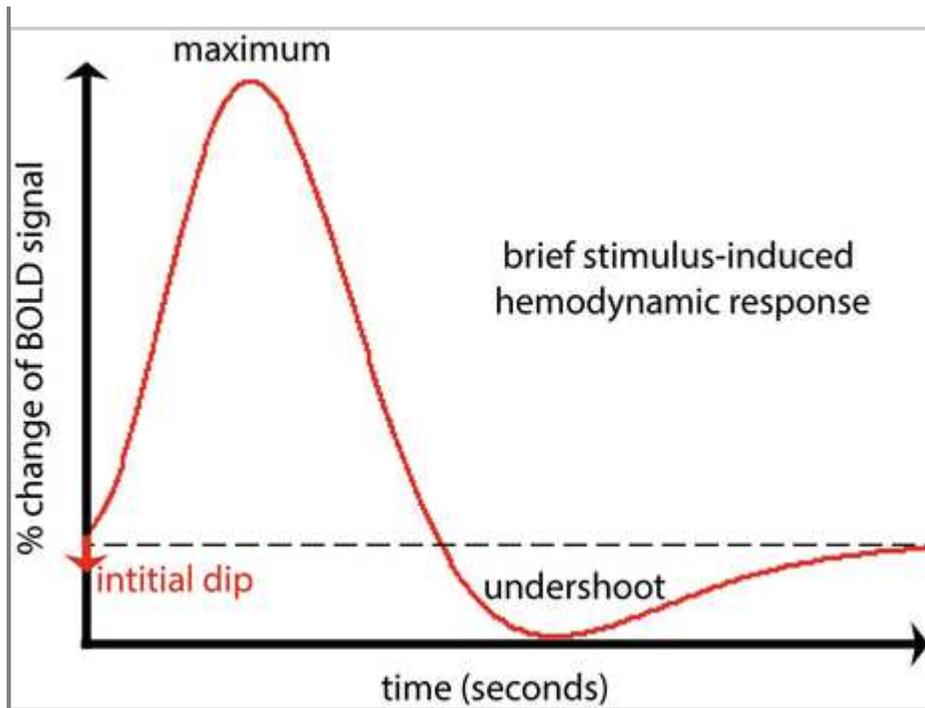
Specific sequences of MRI are applied to encode this hemodynamic response within the $T2^*$ -weighted image such as $T2^*$ -weighted gradient echo (GE) and echo-planar imaging (EPI) with usual spatial and temporal resolutions $2-4 \text{ mm}^3$ and $2-4 \text{ s}$, respectively. BOLD signal detected by EG sequence predominantly derives from venous system. Two reasons may explain this fact. First, upstream, capillaries contain more oxyhemoglobin than veins. Second, in small vessels, erratic moving spins sample all local perivascular magnetic fields, which results in narrowing the phase distribution and consequently in

reduced BOLD signal [12]. However, this effect can be utilized during **spin echo (SE) imaging** since the refocusing 180° RF pulse eliminates only dephasing effects of static magnetic inhomogeneities, but not the phase accumulation of diffusing spins. SE sequence is therefore sensitive to smaller vessels than GE ones.

Phenomenologically, BOLD response, i.e., the percentage of change in MRI signal caused by a very brief stimulation, is characterized by a brief and inconstant initial “dip” followed by the rise of the signal to a peak (4–6 s) due to blood oxygenation increase and followed by slower fall (around 15 s) and finally by a shallow post-stimulus undershoot of variable duration before reaching the baseline (Fig. 2.1). The vascular mechanisms producing the dip and the undershoot remain still a matter of debate. It is noteworthy that BOLD response (20–25 s) lasts longer than neuronal electrical events (several milliseconds) and can exhibit interindividual and interregional variability concerning its time-course. The task-related BOLD increase is usually $<5\%$. Finally, BOLD response can be regarded as a linear function of the underlying neuronal response in first approximation and is correlated with local potential fields which reflect afferent synaptic inputs and postsynaptic dendritic processing rather than spiking [16], although discrepant results found correlation with spiking [11].

Fig. 2.1

Curve representing schematically the **time-varying BOLD response** caused by a brief stimulus applied at time $t = 0$. The dashed line corresponds to the baseline level of BOLD signal



In conclusion, positive BOLD signal mainly reflects intra- and extravascular magnetic field alterations due to focal hyperemia-related, capillaro-venous hyperoxygenation induced by tight neurovascular coupling during neuronal activation. Conversely, negative BOLD signal can be caused by neuronal inhibition. However, regional and stimulus-dependent variations can occur such as inhibitory interneuron-related positive BOLD signal in the cerebellar cortex or striatal negative BOLD signal despite neuronal activation.

CBF-Based Functional Imaging Besides BOLD fMRI, arterial spin labeling (ASL) can also be used for obtaining functional images [17], even in the brain resting state [18]. This MRI sequence usually allows generation of perfusion-weighted images and quantification of absolute perfusion. Put in a nutshell, radio-frequency pulses and gradients applied to the neck invert spin magnetization of hydrogen atoms in arterial blood. In other words, water in the blood is used as endogenous tracer. After a short delay (<5 s) necessary for the labeled spins to reach the studied tissue upstream, two-dimensional EPI images of the brain are acquired. Then, subtraction between paired images with and without spin labeling (control image) enables to suppress static tissue signal and consequently to retain perfusion-related signal. In clinical and research routines, pseudo-continuous ASL is preferentially utilized. Therefore, ASL can detect slow task-related changes of local brain perfusion only due to arteriolar or capillary dilation or contraction. This ASL-based and CBV-related functional mapping improves spatial resolution for localizing accurately activated areas, especially in regions with high susceptibility, and it excludes complex dependency of the signal to other parameters, such as CBV and CMRO₂ modulating the BOLD signal [19]. Moreover, the ability to measure

perfusion permits to evaluate not only interregional signal correlation but also local and global variations of signal amplitude [19, 20]. It is noteworthy that both ASL and BOLD signals can be recorded during a unique dual-echo EPI sequence. The BOLD signal can be extracted from the control (task-related or rest) image.

Brief Survey of Basic fMRI Design and Processing

The aim of the study is to identify brain areas specifically recruited by an experimental task (movement, perception, cognition) performed by the subject inside the MRI machine. Task-related activation must be extracted from the basal metabolism-related activation pattern of the rest of the brain. Two main experimental designs can be applied: *block design* and *event-related design*. In *block design*, the functional run encompasses a regular alternation of “on and off/controlled” blocks. Within each on-block, subjects are requested to carry out the same experimental activity continuously, while, within each off-block, they remain at rest. Block repetition will sample weak BOLD responses thus increasing signal-to-noise ratio, and subtraction of off-blocks from on-blocks will only retain task-specific activated areas. Alternatively, in *event-related design*, discrete stimuli are presented repeatedly but briefly and randomly, and they are interspaced by off epochs of variable duration. Therefore, event-related design is more prone to record transient brain activity than block design more sensitive to sustained brain activity but with a risk of habituation or task anticipation.

Raw functional images must then undergo a complex preprocessing before computing task-related “activation maps.” A preliminary analysis is indeed required since, for instance, raw data are noisy due to several scanner artifacts, head motion, breathing, and heartbeat and encode not only a weak task-related BOLD signal, less than 5% of the total BOLD signal, but also BOLD variations caused by simultaneous but task-unrelated mental activity.

The main steps of *preprocessing of single-subject data* include (1) distortion correction, (2) motion correction and slice realignment, (3) slice-timing correction (correction of interslice delays of acquisition), (4) spatial normalization (realignment of individual anatomical data into a common framework for interindividual comparisons and group analysis), (5) spatial smoothing with a Gaussian kernel (removal of high-frequency noise), and (6) temporal filtering (removal of low-frequency noise such as the scanner drift).

Afterward, statistical modeling and inferences will determine which *voxel’s signal* are significantly and specifically correlated with the experimental design. The most popular method utilized to achieve this goal is the general linear model (GLM). GLM will relate a dependent variable (voxel’s BOLD signal time series from observed data) to one or more independent explanatory variables (voxel’s predicted BOLD signal time-course or perfusion time-course using functional ASL). For example, the predicted BOLD time-course can be modeled by a boxcar, representing alternation of on-off epochs, convolved with a hemodynamic response function. Observed data \mathbf{Y} are, thus, represented as a linear combination of regressors: the predicted neuronal response (design matrix) \mathbf{X} weighted by unknown parameters $\boldsymbol{\beta}$ quantifying effect sizes, and of *Gaussian noise* (unexplained variance) ζ with a null mean and a variance σ^2 , as expressed in the following matrix equation:

$$\mathbf{Y} = [\boldsymbol{\beta}] \cdot \mathbf{X} + \zeta$$

The β values can be estimated by minimizing the sum of squared residuals, i.e., differences between observed and predicted values so that the model best fits the data and, after some mathematical manipulations, can be written as:

$$[\beta] = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y}$$

where \mathbf{X}^t denotes the transposition matrix and provided that $\mathbf{X}^t \mathbf{X}$ is invertible necessitating that \mathbf{X} has a full column rank.

Moreover, GLM requires that the components of the noise ζ remain temporally uncorrelated and share the same variance, which is not always the case. Therefore, after the temporal filtering, prewhitening the data could be applied to remove autocorrelation. Then, a statistical test, such as t-test or F-test, will determine whether the observed data estimated by the set of β parameters and the matrix design may result from random fluctuations of brain activity or of all other null controlled condition (H_0 hypothesis accepted) or not (H_0 rejected: “there is a significant task-related effect”). More formally, the testing relies on contrasts \mathbf{c} defined as a weighted linear combination of β values representing effect of interest:

$$\mathbf{c} = \mathbf{c}^t [\beta]$$

where c denotes the contribution of β to the BOLD signal at a particular TR and in a specific voxel, \mathbf{c} can take -1 , 0 , or 1 values, and then the H_0 and H_1 hypotheses can be expressed by:

$$H_0 : \mathbf{c} = 0 \text{ versus } H_1 : \mathbf{c} > 0$$

If a t-test is applied, the statistics following the t-distribution with (number of rows – number of columns in $\mathbf{X} + 1$) degrees of freedom is expressed by:

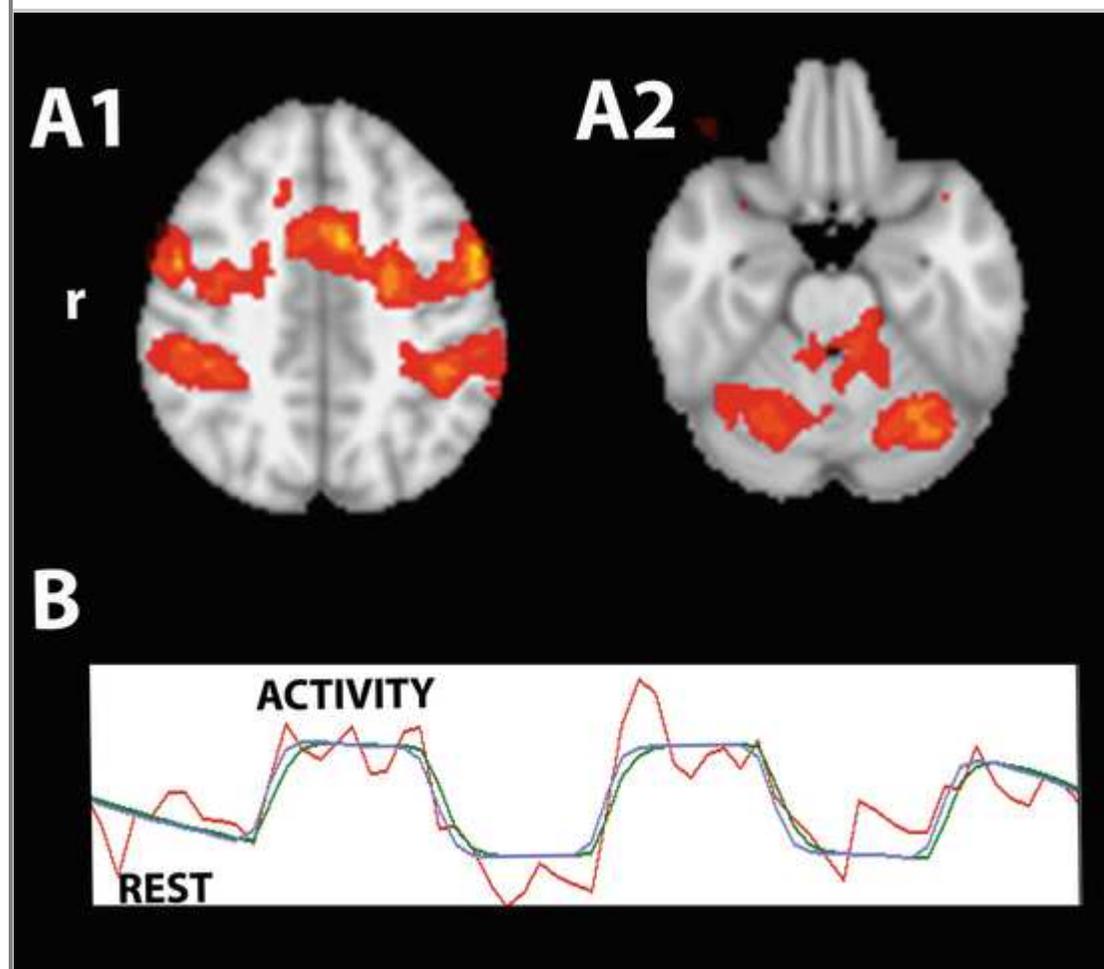
$$\mathbf{t} = \mathbf{c}^T \beta / \sigma \left[\mathbf{c}^T (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{c} \right]^{1/2}$$

Statistical inferences, therefore, require to define the contrast (performance against rest or performance against another performance), the type of level inference (voxel versus cluster), and the statistical threshold (P-value uncorrected or, most often, corrected for multiple comparisons, e.g., $P(T > t) = 0.05$ corrected, where T refers to the random variable following the t-distribution having the same degrees of freedom than the computed t). Multiple comparisons is an important problem. If the probability of false positives is fixed at 0.05 for each test per voxel, then the probability of false positives for all the N studied voxels equals $0.05 \cdot N$ number which can be very high. Therefore, it is absolutely necessary to resort to multiple comparison corrections. Several procedures can be used such as family-wise rate correction (Gaussian random field), for false discovery rate correction controlling the false positives only among the significant voxels. The final results, the statistical parametric map, are visualized on thresholded activation map overlaid on T1 high-resolution anatomical

images, where the voxel statistical significance is represented by a color gradient (Fig. 2.2).

Fig. 2.2

Statistical parametric map of a subject performing alternating phases of rest and activity (motor mental imagery task) during a functional MRI scan. **(a)** Axial slices showing bilateral (pre-)frontoparietal cortical activation (A1) and cerebellar activation (A2). **(b)** The curves represent the detected time-varying BOLD signal (red) and the full model fit versus data obtained by GLM method (blue) (post-processing using analysis of FMRIB Software Library v5.0, Oxford, UK: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT>)



Further higher-level statistics using **single-subject statistical maps** can be performed to create a single-group activation maps and to test intersession or intergroup differences in the brain activation pattern. Generally, as we want to extrapolate information from the limited group of subjects to the whole population, they belong to, random effect analyses are preferentially carried out rather than fixed effect.

Resting-State Functional Connectivity

Definition There exist tight temporal correlations between spontaneous Resting-state functional connectivity

BOLD signal fluctuations at low frequency [0.01–0.1 Hz] and at rest, between spatially remote but functionally related brain areas [21, 22]. The spontaneous brain activity at rest thus results from complex dynamical interactions between well-segregated but partially overlapping, large-scale neural networks involved in motor, perceptual, and cognitive functions [23]. rsfMRI is devoted to identify these resting-state networks (RNs), reminiscent of task-related networks, and their interactions (correlation/decorrelation and anticorrelation) across time. This internally driven brain activity is assumed to reflect, at least, not only the concomitant free conscious experience referred as “mind wandering” but also spontaneous neural events constrained by experience-dependent neuroplastic changes and network structural and/or functional topological architecture. Let’s mention that ASL-based rsfMRI can also be used [18].

Neurophysiological Basis The neurophysiological mechanisms underlying these interregional synchronizations are poorly understood and still a matter of debate. First, interregional functional coherence only partially reflects the anatomical wiring. This fact is illustrated by left and right BA 17 regions, which belong to the same primary visual RN although they are not anatomically interconnected, and by the **dorsal attentional network** and the visual RNs which are functionally separated although they anatomically linked [24]. Notwithstanding, **callosotomy** markedly reduces interhemispheric connectivity. Second, there exist transient and complex correlations between BOLD fluctuations and alpha, beta, gamma, delta, and theta EEG rhythms [25]. For instance, the **default mode network** and the **dorsal attentional network** exhibit positive and negative correlations with beta and alpha rhythms, respectively, whereas the salience networks are active in the gamma range. MEG recordings have also highlighted relationships between BOLD fluctuations and amplitude envelope of band-limited rhythms [26]. Third, in monkey, spontaneous BOLD fluctuations were consistently correlated with neural spiking activity and gamma-range band-limited power [27]. In human, correlations were also found between spontaneous BOLD and the slow cortical potentials. Fourth, several other putative factors have been assumed to shape BOLD spontaneous fluctuations [24, 28]: **cellular metabolism** (redox variation associated with energetic metabolism), biophysics (up-down state transition of the membrane voltage, subthreshold oscillations, microstates), biochemistry (quantal exocytosis), network constrains (time delay, noisy transmission, nonlinear attractor dynamics [29]), anatomical connections (thalamus), or neuromodulation. Therefore, spontaneous BOLD fluctuations seem to be spatially and temporally multiscale organized from the cell to the network.

MRI Sequence Inside the scan and during the whole experiment, subjects are requested to remain still and eyes closed and to focus their attention upon nothing in particular. Then, the same EPI sequences than in classical fMRI are applied to measure the resting-state brain activity. For example, around 200 volumes covering the whole brain are acquired with a TR = 2 s. Utilization of accelerated sequences, such as multiband (simultaneous acquisition of multiple slices), can notably diminish acquisition time and TR value (<1 s) and/or to increase spatial resolution. It is noteworthy that the raw data will be preprocessed as in GLM method.

Statistical Analysis The two main techniques for performing functional connectivity analysis are **region-of-interest (ROI) correlations** and **independent component analysis (ICA)** [30]. The first method focusing on a specific brain area is based on calculating correlation (Pearson score) of the BOLD signal between a predefined ROI and the rest of the brain. Alternatively, ICA is a multivariate, data-driven method allowing the identification of multiple coexisting whole-brain networks and subnetworks by voxel-to-voxel analyses. To achieve this goal, ICA solves the *blind source separation problem*, which requires distinction between a set of unknown sources underlying the observed data. ICA assumes that these sources are statistically independent, non-Gaussianly distributed, and linearly mixed. For functional imaging and more formally, spatial ICA seeks to separate all the *hidden* distinct neural networks \mathbf{X} and, in probabilistic ICA model, possible source of normal noise ξ contributing altogether and simultaneously to the observable whole-brain BOLD signal \mathbf{Y} [31]:

$$\mathbf{Y} = \mathbf{M} \cdot \mathbf{X} (+\xi)$$

where \mathbf{M} denotes the *unknown mixing* matrix. Sometimes, the raw data have been beforehand prewhitened by principal component analysis in order to reduce the dimensionality of data spaces and to remove Gaussian noise. ICA algorithm, such as maximization of negentropy or maximum likelihood, for example, then aims to determine the *unmixing* matrix in order to recover the unknown sources \mathbf{X} from \mathbf{Y} . This algorithm computes a set of statistically independent components using, for instance, maximization of non-Gaussianity (using higher-order statistics). Each neural network and artifactual sources \mathbf{X} correspond to a specific ICA component (spatial maps) and are associated with a specific BOLD time-course. The number of components remains a free parameter which can be fixed a priori (around 15–20) or estimated by the algorithm (around 30–70): the lower the number (around 20), the more robust and reproducible results. Excessive number of components might split networks into subsystems.

Group and intergroup analyses can also be conducted by ICA [31, 32]. For instance, temporal concatenation supposes common spatial maps associated with unique time-courses across subjects. The subject-specific maps and time-course can, then, be back-reconstructed on the basis of the group-level components. Finally, several algorithms enable intersession or intergroup comparisons such as dual regression and randomization. It is worth emphasizing that algorithmic results can vary in function of the population size and because of its probabilistic nature.

Finally, other methods can be applied in the “resting” brain such as:

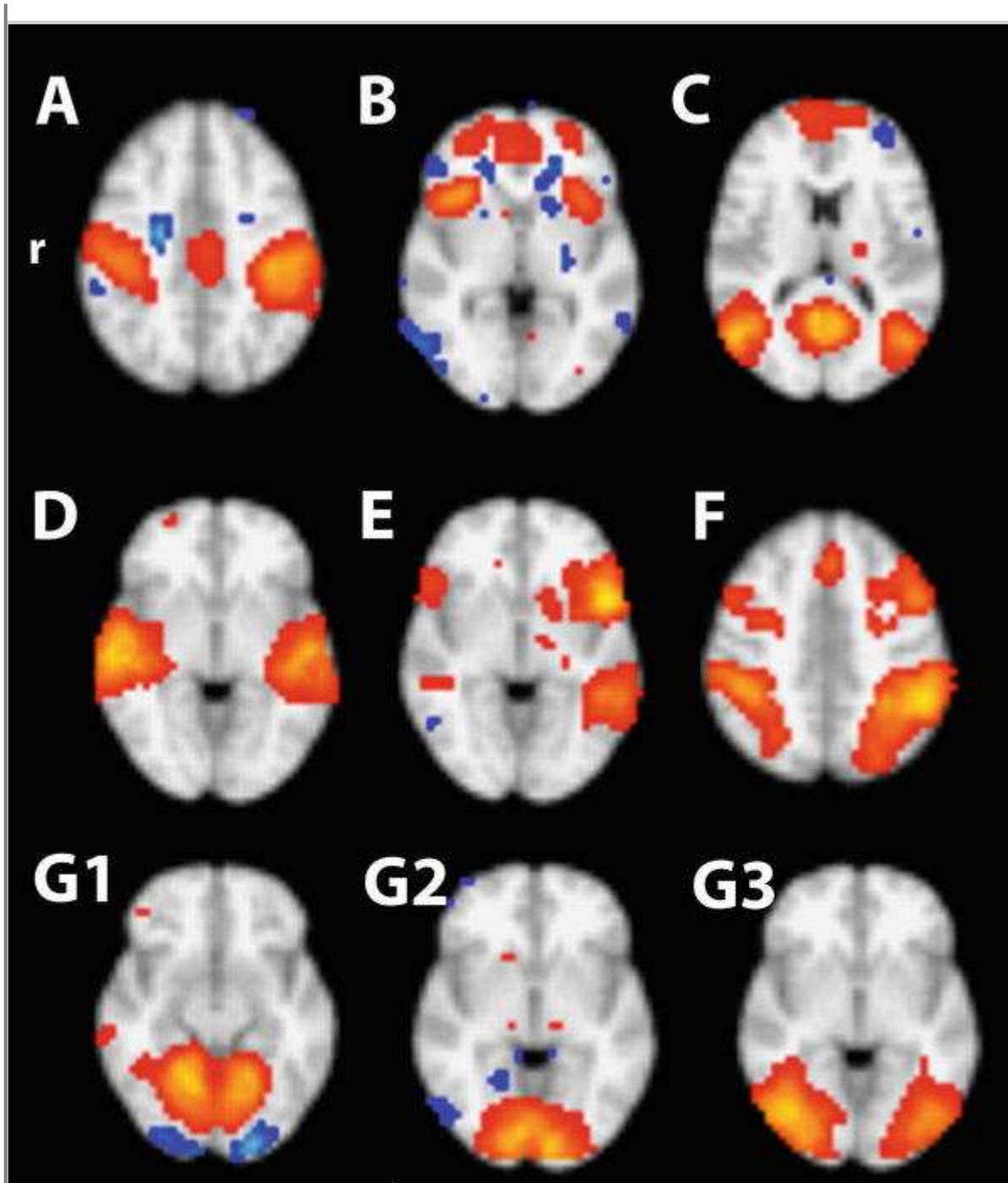
- Detection of (fractional) “amplitude of low-frequency (0.01–0.1 Hz) fluctuations (ALFF)” calculating the intra-voxel power spectrum of the BOLD signal [33]
- “Regional homogeneity” (ReHo) approach measuring synchronization between time-course of BOLD signal of neighboring voxels using the Kendall’s coefficient of concordance [34]

However, the resulting functional maps reflect local processing and not large-scale functional circuits.

Resting-State Network Identification The main step after having obtained the ICA components relies on the selection of components representing the true neural networks. Some criterion can be usefully applied and automated. First, only components whose frequency is comprised between 0.01 and 0.1 Hz or, at least, whose more than 50% of the spectrum remains within this interval, must be retained. Second, noisy components must be well-identified (structured noise, head and ocular motion, breathing, cardiac beating, scanner drift, etc.) and removed. Several typical aspects of noise must be kept in mind: *structural noise* like ring pattern of pixel around the encephalon due to head motion, diffuse spotty pattern over the brain, large or tighter bilateral frontal and orbitofrontal clusters with “spike” or “sawtooth” motif appearing in the corresponding time-course, respectively, clusters located in the ventricles, and *improper anatomical regions* like white matter, cisterns (especially around the brainstem), superior longitudinal sinus, or eyes [35]. This can be achieved by (operator-dependent) visual inspection or by automated classification algorithms. It is possible to take advantage of the extraction of noisy components in single-subject components, to denoise the raw data, what will improve further group analysis. Third, most of the RNs have been listed and can be easily recognized according to anatomical localization of the network nodes [36, 37] (Fig. 2.3):

Fig. 2.3

Axial slices of the brain showing fMRI resting-state networks computed by independent component analysis. **(a)** Motor network. **(b)** Salience network. **(c)** Default mode network. **(d)** Auditory network. **(e)** Language network. **(f)** Dorsal attentional network. **(g)** Medial (G1), caudal (G2), and lateral (G3) visual networks. **(h)** Right (H1) and left (H2) executive networks. r: right



1. The **default mode network (DMN)** (bilateral medial prefrontal, posterior cingulate, retrosplenial, precuneal, lateral inferior parietal, and parahippocampal cortices, cerebellar amygdalae, and cerebellar lobule VII) involved in consciousness, unconstrained cognition (“mind wandering”), episodic memory, emotional processing, and self-projection in future or past
2. The **left executive control network (LECN)** (left dorsolateral prefrontal and superior parietal cortices, contralateral cerebellar lobules HVI and HVII) involved in syntactic cognition and language
3. The **right executive control network (RECN)** (right dorsolateral prefrontal and superior parietal cortices, contralateral cerebellar lobules HVI and HVII) involved in holistic and visuospatial cognition
4. The **dorsal attentional networks (DAN)** (bilateral intraparietal sulcus, prefrontal, precentral cortices) for top-down focus of attention and goal-directed cognition
5. The **salience network (SN)** (prefrontal operculum, anterior insula, dorsal anterior cingulate, hypothalamus, and cerebellar lobules HVI and HVII) involved in interoception, emotion, and attentional reallocation based on stimuli salience
6. The **sensorimotor network** (bilateral S1-M1, supplementary motor area, cerebellar anterior lobe and lobule VIII)
7. The **medial primary visual cortex**
8. The **polar visual cortex**
9. The lateral (extrastriate) visual cortex
10. The **auditory network** (superior temporal, insular, and postcentral cortices)
11. The **linguistic network** (superior and medial temporal, inferior frontal, and angular cortices)
12. The striatum
13. The cerebellum

SN, LECN, RECN, and DAN are grouped in a **generic external attention system (GEAS)** in opposition to the DMN implicated in self-referential thinking.

GEAS and DMN appear to work most often in counter-phase. More precisely, DMN is anticorrelated to the DAN but can be flexibly correlated, for example, during goal-directed mental simulation or anticorrelated to ECN. The dorsolateral prefrontal region of the ECN can inhibit the medial prefrontal part of the DMN [38]. It is suggested that the SN may exert a role in switching between these two internally and externally oriented systems in response to behaviorally relevant salient external stimuli. Moreover, a seed-based analysis has characterized a ventral right-lateralized attention system encompassing the ventral frontal cortex and the temporoparietal junction [39]. It turns out that the switching mechanism would rely on antagonistic connectivity between the right insula and the dorsal posterior cingulate part of the DMN. At rest, only part of these networks such as DMN and SN are currently active. Finally, hierarchical clustering using, for instance, partial correlation matrix could contribute to distinguish and rank RN and their subnetworks.

Time-Dependent Properties of Resting-State Networks These networks evolve across time. Although RNs are robustly detected in fetus, children, asleep or anesthetized subjects, and animals (monkeys, rats, and mice), the functional connectivity is age-dependent and context-dependent as it varies in function of several factors such as open versus closed eyes, mood, drowsiness, mental content, and previous cognitive effort. First, age-dependent changes mainly concern intra- and inter-network reorganization of connectivity. In fetus, cortico (M1)-cerebellar, thalamocortical, interhemispherical coherences have been observed [40]. Moreover, a proto-DMN interconnecting medial prefrontal and precuneal cortices has also been identified after 35 weeks. In children, bilateral sensorimotor, auditory, primary visual anterior prefrontal, and medial and lateral parieto-cerebellar networks are the first RN to be discernible by age 1. A bit later, DMN, SN, and executive control networks can be visualized by age 2 but will undergo further gradual maturation processes including increasing within-network coherence, eliminating some distant interconnection, and developing long-range connectivity. Motor and perceptual RNs are mature earlier than higher cognitive circuits. In older people, DMN shows the same global connectivity than in young adults but with weaker connectivity between prefrontal and ventral precuneal cortices and stronger correlation in the dorsal precuneal circuit [41]. Second, functional connectivity is not only influenced by specific mental states, as mentioned above, but also by practice of motor, perceptual, and cognitive tasks. In particular, more or less transient synchronization may occur between performance-related brain areas and ECN. DMN exhibits task-induced deactivation and, more often, subsequent anticorrelation with networks subserving overt sensorimotor, attentional, or cognitive activity like stimulus-conscious perception: the weaker the anticorrelation, the poorer the performance likely due to interference between mind wandering and task-related attention. However, some nodes of the DMN, such as the right posterior cingulate cortex, can collaborate with ECN during memory recollection and improve rapidity of the task. Therefore, competition or cooperation between RN and strengthening interconnections may reflect transient “scaffolding” connectivity (attention to the task) and neuroplastic changes caused by practice (off-line consolidation, memory trace, and performance ability), respectively [42]. Third, classical studies provide a set of apparent stationary RN during the experiment. However, more refined post-processing using a sliding time window has demonstrated time-varying interactions of RN subnetworks and between RNs [43]. Moreover, a component related to drowsiness and light sleep has been characterized including breakdown of DMN connectivity, reduced thalamocortical coherence, and increased subcortical connectivity. In summary, the whole set of RNs is common to all subjects; their complex pattern of interactions shaped by experience-dependent plasticity may reflect individual idiosyncratic functional and structural connectivity and, consequently, personal cognitive profile.

Clinical Application Functional connectivity enables not only to explore normal brain functioning but also brain functional impairments in neurology (stroke, multiple sclerosis, pain, migraine, amnesia, spinocerebellar ataxia, Parkinson’s disease, epilepsy, altered consciousness like vegetative state and coma), neuropsychology (sleep, aging), pharmacology (treatment evaluation, drug abuse), neuro-ophthalmology (glaucoma, visual deprivation),

and psychiatry (attention-deficit/hyperactivity disorder, obsessive and compulsive disorder, Tourette syndrome, post-traumatic stress disorder, depression, neglect, autism, schizophrenia, bipolar trouble, mild cognitive impairment, dementia), as well as potentially after neurostimulation (TMS, TDS) [44, 45]. It can also be used for presurgical planning and postsurgical follow-up helping to localize precisely (sensorimotor or language) functional areas. This method offers important clinical possibilities. It can be applied to sedated, unconscious, or uncooperative patient, especially patients unable to perform task-based fMRI. It permits to screen a wide range of functional networks either in an exploratory mode (ICA) or in focal seed-based mode and to determine qualitative (anatomical) and quantitative (inter-areal correlation values) RL alterations in comparison with matched normal subjects. These data can provide markers for diagnosis and prognosis and anatomico-clinical correlations.

Effective Connectivity

Effective connectivity can be defined as the influence that one neural system exerts over another, either at a synaptic level (synaptic efficacy) or a cortical level [46]. This hypothesis-driven approach emphasizes that determining effective connectivity requires a causal model of the interactions between the elements of the neural system of interest. Causal relationships are not inferred from the data but are assumed a priori. If fMRI and rsfMRI identify statistical covariation across functionally related neural nodes, they do not provide any information about direct causal influence between these nodes: “does activated node A directly cause or modulate activation of node B? or, alternatively, “does a third node D synchronize independent nodes A and B?”. Effective connectivity tries to address these important points.

First, tractography may delineate which nodes are anatomically interconnected – however within the limits of spatial resolution and fiber crossing/convergence detection – since functional interconnection presupposes axonal interconnection. Second, several complementary methods inferring internodal causality based upon fMRI data have been developed such as dynamic causal modeling (DCM), structural equation modeling (SEM), time series analyses, and brain perturbation (with transcranial magnetic stimulation, for instance) [47].

DCM roots in the hypothesis that brain can be regarded as a deterministic nonlinear dynamic system transforming inputs into outputs and taking into account modulations [48]; DCM models the effects of experimental, external, and modulatory inputs on network dynamics. Briefly, the first step of this method consists in building a biologically plausible neuronal model of interacting brain regions detected by fMRI. Each of these neural regions is characterized by its state z , representing its amount of activation, and its coupling with its interconnected regions. The time evolution of z is thus described by a nonlinear differential equation: $\dot{z} = dz/dt = [A + \sum B^i \cdot u_i + \sum D^j \cdot z_j] z + C \cdot u$ which depends on the current state z , the intrinsic internodal coupling (A matrix), the transient changes in intrinsic coupling due to j th internal input (B matrix), the direct action of external inputs u upon the region (C matrix), and the D matrix (nonlinear state equation) which encodes how the n regions gate connections in the system. However, only the hemodynamic response underlied by the neuronal activity is measured using fMRI. Therefore, this neuronal model must be complemented by a hemodynamic model biophysically plausible in terms of vasodilatory signal which induces increased blood flow. This results in local changes in blood volume and deoxyhemoglobin (dHb) from which the BOLD signal can be predicted. The aim of the DCM is to estimate neural (A , B , C , D) and hemodynamic parameters (flow, volume, etc.) such that the modeled and measured BOLD signals are maximally similar. Afterward, the parameters of the neuronal model can be estimated using Bayesian framework. The oversimplification of the anatomical/connectivity architecture of the neuronal networks constitutes one of the main limitations of this

method.

SEM is a multivariate method which assigns specific strength (or *path coefficient*) to anatomical connections linking functionally related brain areas. Unlike DCM, SEM assumes instantaneous (non-dynamic) interactions and the inputs stochastic and unknown. So, SEM models can be considered to be “static” as they model instantaneous interactions between regions and ignore the influence of previous states on current responses. Path coefficients are computed by minimizing the covariance matrix of the modeled (theoretical) network and the covariance matrix of the observed data (BOLD response) – the amount of statistical dependency between the time series of the BOLD signal extracted from pairwise associated brain regions. Here again the relevance of **SEM** depends on the anatomical accuracy of the model and on the linear assumption of the change in path coefficients and estimates of path coefficients limited to the hemodynamic level.

Coherence analysis studies the correlation in the frequency domain, using fast **Fourier transform**, between BOLD responses of interconnected brain areas. This spectral method can also determine the causal directionality by calculating the sign of the slope of the phase spectrum (for low frequencies), and the inter-areal time lag by the absolute value of the slope. Alternatively, **Granger causality method** uses a linear autoregressive prediction claiming that current state of activation of region A, at t , is predicted by a linear combination of its previous states measured at $(t-nTR)$. It further hypothesizes that if activation of region A causes activation of region B, then adding the autoregressive model of A to autoregressive model of B will improve the predictability of B activation in comparison to the autoregressive model of B alone. Then, the difference between observed data and the predictive model, estimated by the variance, can be used as a measure of adequacy of the model and, up to some algebraic transformation, of inter-areal causation.

Coherence analysis and **Granger causality** suffer from several common and specific weaknesses which prevent their utilization in fMRI [49]. The main problem relies on the difference between the duration of neuronal processing on the order of milliseconds, the rate sampling of BOLD data on the order of seconds, and the total duration of BOLD response (20–30 s). Therefore, the DCM method seems the most appropriate one to characterize causal directionality between brain regions.

Finally, one last complementary method to test potential causal influence between two brain areas consists in actively perturbing the functioning of a given brain area. Magnetic or electric transcranial stimulation can be used to activate or inhibit a specific brain region and to study the influence of this perturbation on the activation of interconnected areas.

Graph Theory

Graph theory

Graph theory

has already proven to be applicable to a considerable diversity of complex systems, including markets, ecosystems, computer circuits, and gene-gene interactomes. **Complex network theory** is particularly appealing when applied to the study of clinical neuroscience, where many cognitive and emotional disorders have been characterized as **disconnectivity syndromes**, as indicated by abnormal phenotypic profiles of anatomical and/or functional connectivity

between brain regions. For example, in schizophrenia, a profound disconnection between frontal and temporal cortices has been suggested to characterize the brain; in contrast, people with autism may display a complex pattern of **hyperconnectivity** within frontal cortices but **hypoconnectivity** between the frontal cortex and the rest of the brain [49].

Graph theoretical analysis is potentially applicable to any scale, modality, or **volume of neuroscientific data** (e.g., diffusion tensor imaging (DTI), diffusion spectrum imaging (DSI), EEG, MEG, cortical thickness, resting state, task-related fMRI) providing new measures of human brain organization in vivo. It is conceptually easier to link the brain graphs derived from these different data types to each other than it would be if each imaging dataset were described in terms of some modality-specific measure of association between regions, e.g., tractographic connection probabilities from DSI or correlations between regional fMRI time series. Facilitation between-modularity translation of results can be important for methodological cross-validation and, more fundamentally, for informing our understanding of how functional networks might interact with the substrate of a relatively static structural network. For example, fMRI and DTI brain graphs consistently demonstrate some common global topological properties allowing a high efficiency at different spatial and temporal scales with a very low wiring and energy cost [50].

We summarize below few of the **topological measures** that have been most extensively investigated in the neuroimaging literature to date [50, 51]. Any **complex dynamical network** will be mathematically described as graphs that represent a set of n nodes or vertices associated with k connections or edges/lines between them.

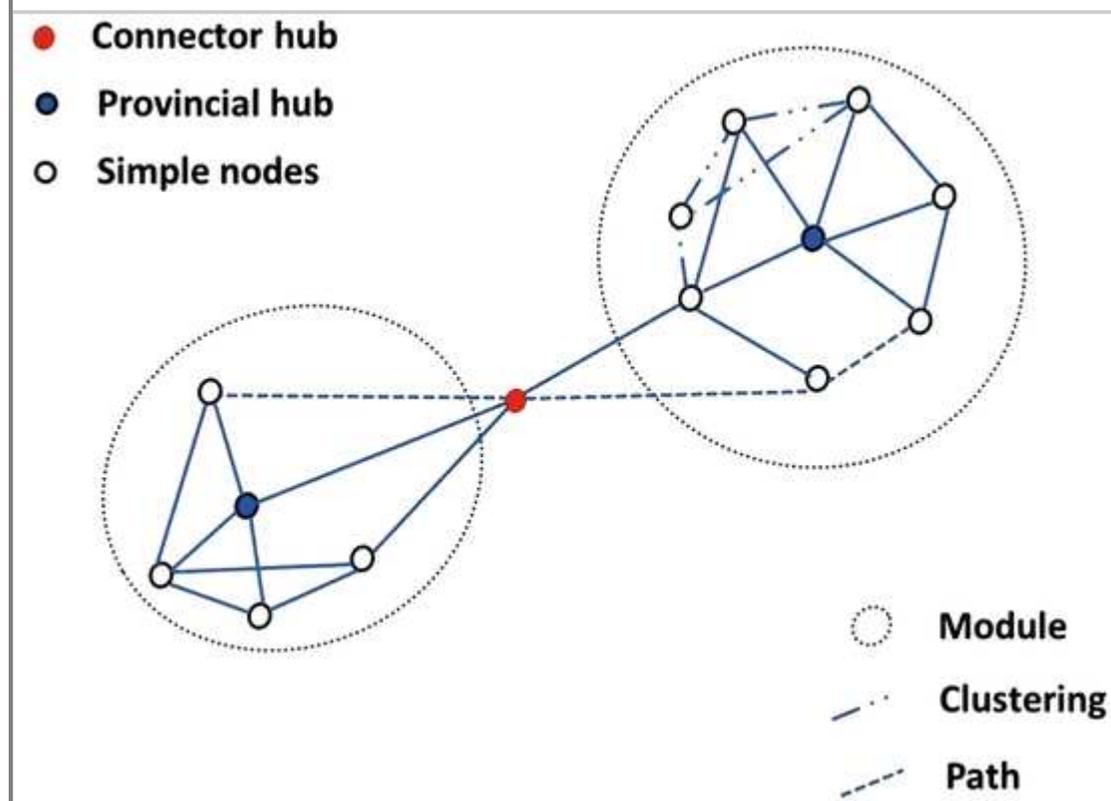
The “small-world” network, the first mathematical model originally described in social networks, combines high levels of local clustering among nodes of the network and short paths that globally link all nodes of the networks. Small-world organization (many short path lengths, few long path lengths) is intermediate between the random networks (all nodes are related randomly with same probability for short- and long-distance connections) and the regular networks or lattices where all nodes are only related to their nearest neighbors (many short path lengths, no long path length). Path length is the minimum number of edges that must be traversed to go from one node to another. **Random and complex networks** have short mean path lengths (high global efficiency of parallel information transfer), whereas regular lattices have long mean path lengths. Efficiency is inversely related to path length but is numerically easier to use to estimate topological distances between elements of disconnected graphs. *Small-world* and *efficiency* use two metrics: the *clustering coefficient* (C) which is a measurement of the efficiency of local connectivity and the *path length* (L) which is a simplified measurement of the global efficiency of information transfer on the network. These two metrics enable to define the small-world properties, in which the network exhibits a clustering coefficient greater than the clustering coefficient of a random network and a path length about the same as the path length of a random network. The small-world scalar dependent on the calculation of a path length can be troublesome for networks that contain one or more disconnected nodes. The path length of a disconnected node is infinity; it cannot transfer information to any other node on the network. In a complementary formalism, the **global efficiency** is introduced as an alternative metric of global integration that is inversely proportional to the characteristic path length of the network, thus allowing computation of a finite value for graphs with disconnected nodes.

Many complex networks consist of a number of modules [50]. There are various algorithms that estimate the *modularity* of a network, many of them based on hierarchical clustering. Each module contains several densely interconnected nodes, and there are relatively few connections between nodes in different

modules. A brain graph can be subdivided or partitioned into subsets or modules of nodes. In general, the aim is to find the partition that maximizes the ratio of intramodular to intermodular edges. The nodes in any module will be more densely connected to each other than to nodes in other modules. The *intramodular degree* is a measure of the number of connections a node makes with other nodes in the same module. The *participation coefficient* is a measure of the ratio of intramodular connectivity to intermodular connectivity for each node. Finally, these and related metrics can be used to define nodes as “connector hubs” with high intermodular connections or “provincial hubs” with high intramodular connections. Therefore, hubs can be described in terms of their roles in a community (modular) structure (cf. Fig. 2.4).

Fig. 2.4

Representation of network with modular organization. Provincial hubs are connected mainly to nodes in their own modules, whereas connector hubs are connected to nodes in other modules



Degree and degree distribution indicate the likely presence of network hubs or highly connected nodes. The simplest topological measure is the *degree* of a node which is defined as the number of edges emanating from that node. Degree sometimes called *degree centrality* has been used to discriminate between nodes in the system that are well connected, i.e., so-called *hubs*, and nodes that are less well connected or non-hubs. Due to their relatively increased connectivity, high-degree nodes are likely to play an important role in the system's dynamics. The probability distribution for nodal degree is the *degree distribution* of the network. Brain graphs generally have heterogeneous or broad-scale degree distributions, meaning that the probability of a highly connected hub is higher than in a comparable random network. Most studies have found that an exponentially truncated power law is the best form of degree distribution to fit to networks based on functional and structural MRI data.

Connectivity degree is one of the most basic and important measures of network analysis. The degree K_i of a node i is defined as the number of connections to that node. Nodes with a high degree are interacting with many other nodes in the network. The degree K of a graph is the average of the degree of all N nodes in the graph G :

$$K = \frac{1}{N} \sum_{i \in G} K_i$$

The *connection strength* is a measure quantifying how closely network nodes are connected in terms of showing a relationship in their time-course of activation. The overall connection strength S_i is calculated as:

$$S_i = \sum_{i \neq j \in G} Kr_{i,j}$$

Measures of functional segregation quantify the presence of functionally related, densely interconnected groups of brain regions, known as clusters within the network. The local (nodal) clustering coefficient C_i is defined as the number of existing connections among the node's neighbors divided by all their possible connections:

$$C_i = \frac{E_i}{K_i(K_i - 1)/2}$$

where E_i is the number of existing connections among the node's neighbors.

The clustering coefficient of a network is the average of the clustering coefficient of all nodes:

$$C = \frac{1}{N} \sum_{i \in G} C_i$$

in which C quantifies the extent of local connectivity of the network.

Measures of functional integration characterize the ability to rapidly combine specialized information from distributed brain regions and are commonly based on the concept of a path, with shorter paths implying stronger potential for integration. The mean path length L_i of a node i is:

$$L_i = \frac{1}{N-1} \sum_{j \in G, j \neq i} L_{i,j}$$

in which $L_{i,j}$ is the smallest number of edges that must be traversed to make a connection between node i and node j . The average inverse shortest path length is a related measure known as global efficiency of a network.

In terms of *clinical application*, Redcay et al. [52], to assess whole-brain network properties in adolescents with autism, collected *resting-state functional connectivity MRI (rs-fcMRI)* data from neurotypical adolescents (NT) and adolescents with *autism spectrum disorder (ASD)*. Task-independent studies of the resting brain provide a window with which to examine intrinsic functional network organization. Recent findings suggest connectivity differences in autism with evidence for both hypo- and hyperconnectivity for short- and long-distance connections depending partly on age of the participants.

In this study, the authors used graph theory metrics on *rs-fcMRI* data with 34 regions of interest (i.e., nodes) that encompass four different functionally defined networks: *cingulo-opercular, cerebellar, frontoparietal, and DMN (default mode network)*. These networks were selected because previous research with these same networks has demonstrated a developmental pattern of progressive increases in long-distance connectivity between nodes of the same network and concurrent decreases in connectivity between anatomically proximal nodes of distinct networks. In addition, functions associated with these networks have all been implicated in autism. Thus, examining these networks allows for a more rigorous test of the hypothesis of reduced long-distance and increased local connectivity in autism, across multiple networks that support varied functions.

As mentioned above, graph theory methods can examine the *topological properties* of each region within the context of all other regions of interest, including measures of the integration (global efficiency, path length), segregation (local efficiency, clustering coefficient), and centrality (or betweenness centrality) of networks. These metrics provide a more robust test of the theory of reduced long-distance and increased local connectivity by testing differences in measures of whole-brain network integration and segregation.

In the Redcay's work [52], data were analyzed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) and CONN functional connectivity toolbox (<https://www.nitrc.org/projects/conn>). The *unweighted ROI-to-ROI correlation matrices* were first thresholded at a cost value of $k = 0.15$. Cost is a measure

of the proportion of connections for each ROI in relation to all connections in the network. When cost is equated across participants, direct comparisons across groups of network property differences can be made. Small-world properties are observed in the range of costs $0.05 < k < 0.34$, where global efficiency is greater than that of a lattice graph and local efficiency is greater than that of random graph. A cost threshold of 0.15 has been demonstrated to provide a high degree of reliability when comparing session-specific estimates of graph theoretical measures across repeated runs or sessions (e.g., global efficiency $r = 0.95$, local efficiency $r = 0.9$). The specific measures of interest were those of integration (global efficiency), segregation (local efficiency), and centrality (betweenness centrality). Between-group t-tests were used to compare networks measures between groups with a FDR correction of $p < 0.05$.

Global efficiency is calculated as the average of the inverse of the shortest path length between each ROI (or node) and all other ROIs. The shortest path length is defined as the fewest number of connections (or correlations) between two nodes. Thus, a network with high global efficiency would be one in which nodes are highly integrated, so the path between nodes is consistently short. This measure (with cost kept constant) can be thought of as reflecting global, long-distance connections within the brain. **Local efficiency** is calculated as the average inverse of the shortest path length between the neighbors of any given node (or ROI). In other words, local efficiency measures the extent to which nodes are part of a cluster of locally, interconnected nodes. **Betweenness centrality** (measure of centrality) measures the fraction of all shortest path lengths in a network that pass through a given node. Thus, if a node is directly connected to many other nodes in the network, it will have a shorter overall path length and function as a hub within and between networks.

Contrary to their hypotheses, Redcay et al. found no differences in measures of global or local efficiency. Only betweenness centrality, which indicates the degree to which a seed (or node) functions as a hub within and between networks, was significantly different between groups, and it was greater for participants with autism for the right lateral parietal (RLatP) seed of the DMN only. Using RLatP region as a seed region, authors demonstrate significantly functional connectivity in the ASD than NT group within anterior medial prefrontal cortex (aMPFC) using a few cluster correction of $p < 0.05$; NT group showed higher connectivity between the RLatP seed and cerebellar tonsils (a region previously associated with the DMN). The author suggests that the higher betweenness centrality in ASD may be due to greater long-distance connectivity within the DMN (right lateral parietal-anterior medial prefrontal cortex). They conclude that greater connectivity within right parietal cortex could indicate less functional specialization of this region in ASD.

Machine Learning in Neuroimaging

Machine learning (ML)

Machine learning (ML)

algorithms developed in artificial intelligence and in statistics are increasingly applied to neuroimaging data for automatic and adaptive detection, classification, and prediction of complex, brain structural and functional patterns [53, 54]. From a neurological and psychiatric standpoint, ML can be potentially used for image processing (e.g., anatomical segmentation, image registration), lesion detection, radiomics, diagnosis, and prognostics, while such methods would allow to decode/encode brain states, for instance, in the field of fundamental neurosciences.

As structural and functional brain images contain a huge number of data (voxels), often noisy and spatially correlated, robust multivariate methods applicable to high dimensional data space are requested to uncover and to interpret complex or subtle latent, highly informative, structure. ML algorithms

are, therefore, particularly well suited to fulfill this task and, contrary to univariate investigations, offer the additional possibility to draw inference at the single-subject level. Notwithstanding, ML necessitates large samples of training data, such as MRI scans, so as to decrease the risk of overfitting (inability to generalize from the data used for training to new data) [54].

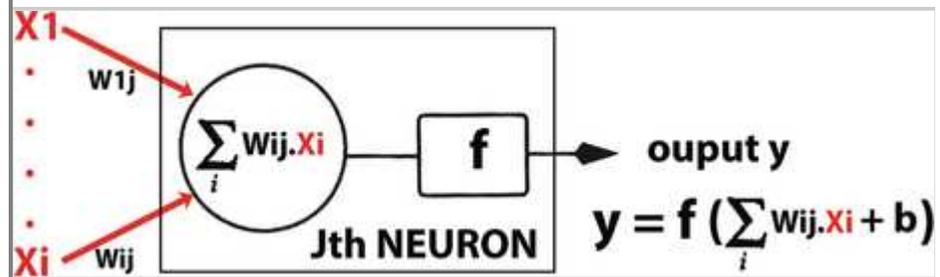
ML Brief Overview Put in a nutshell, ML is trained with data so that the algorithm progressively learns how to identify specific structured patterns “buried” within these same data and how to predict the characteristics of new data. For example, ML can be utilized as a classifier segmenting and categorizing discrete data into groups and ascribing new data to their most likely belonging class. The ML algorithm can be trained in either a supervised or in a (semi-)unsupervised manner.

Supervised ML, such as a specific type of perceptron called support vector machine (SVM) or multilayered perceptron, infers from the training data including both the input data \mathbf{x} and the desired output data \mathbf{y} , the function $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, which enables to predict the features of new data \mathbf{z} : $\mathbf{f}(\mathbf{z}) = \mathbf{y}$. When the input data are continuous, this function is tantamount to a regression. Conversely, unsupervised ML strives to determine the probability distribution of the inputs \mathbf{x} , based only on the training dataset. One important example of unsupervised ML consists in recent deep-learning algorithms [55], which construct automatically hierarchical representations of the input intrinsic organization, from simpler to more complex and abstract concepts, in contrast to supervised ML which needs relevant feature extraction by an instructor. Unsupervised, data-driven ML training has consequently the benefit of (partially) getting rid of any operator dependency.

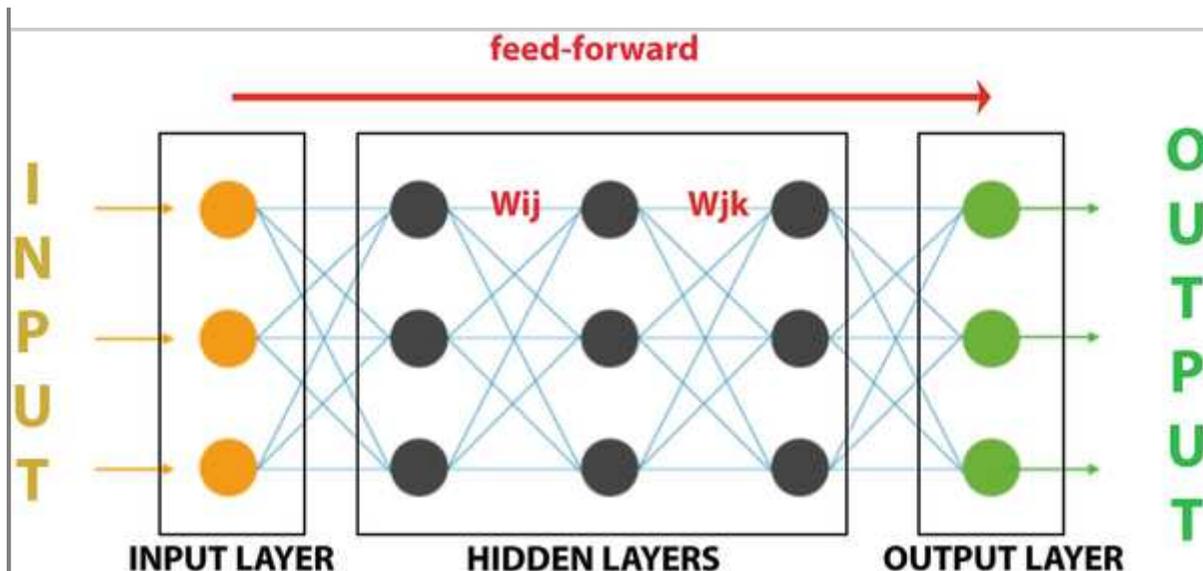
Perceptrons are particular implementation of artificial neural networks (ANN). ANNs are composed of a set of successive connected layers, each layer containing a given number of artificial neurons. Artificial neurons are formal automata which transform the sum of their weighted $\mathbf{w}\cdot\mathbf{x}$ inputs into an output \mathbf{y} : $\mathbf{y} = \mathbf{f}(\sum \mathbf{w}\cdot\mathbf{x} - \boldsymbol{\theta})$, where \mathbf{f} and $\boldsymbol{\theta}$ denote the activation function such as a logistic sigmoid function and an activation threshold (or bias), respectively (Fig. 2.5). Artificial neurons are grouped into “juxtaposed” layers: an input layer, one or several hidden layers, and an output layer. For instance, the output layer could indicate the probability of an input to fall into a given class. In a feed-forward ANN, each neuron is connected to all neurons belonging to the next layer, so that activation propagates feed-forwardly from the input layer up to the output layer after having undergone successive nonlinear transformations within the hidden layers (Fig. 2.6). The number of hidden layer determines the “depth” of the network. Of course, a lot of ANN architectures exist in relation to the number of artificial neurons, their activation function, the network “depth,” and the connectivity pattern (feed-forward, recurrent, or full-connected networks). Perceptrons are trained by data inputs \mathbf{x} , and the corresponding computed output \mathbf{y} is compared with the desired output \mathbf{o} . The difference between these two outputs ($\mathbf{o} - \mathbf{y}$), called error, serves to correct backward and step-by-step the weights of the connections between neurons from the output layer down to the input layer. Stochastic gradient descent algorithm is used as the procedure of error minimization which determines how much to adjust iteratively the weights in order to lower the error. A bit more formally, a loss function measuring the error can be defined as: $\mathbf{E} = \frac{1}{2} \sum (\mathbf{o} - \mathbf{y})^2$. Then, minimizing E requires to find the steepest (ideally, the global) minima of the derivative of E with respect to the weights \mathbf{w} : $\min [\partial\mathbf{E}/\partial\mathbf{w}]$ and to iteratively update these weights, using the gradient descent: $\mathbf{w}(\mathbf{t} + 1) = \mathbf{w}(\mathbf{t}) - \boldsymbol{\eta} \cdot \partial\mathbf{E}/\partial\mathbf{w}$, where $\boldsymbol{\eta}$ designates the learning rate. **The main risk is to fall into a local minima. Stochastic gradient descent algorithms resort to a certain amount of noise injected in the descent gradient process in order to prevent this potential risk.**

Fig. 2.5

Artificial neuron. The weighted sum of the n inputs i : $\sum w_{ij} \cdot x_i$ is transformed by the nonlinear activation function into the j th neuron output: $y = f(\sum w_{ij} \cdot x_i + b)$, which is sent to all next connected neuron. Note that b denotes a bias or an activation threshold

**Fig. 2.6**

Multilayer perceptron. It is a feed-forward artificial neural network composed of an input layer, several successive hidden layers, and an output layer. Each layer is composed of artificial neurons, the neurons of the hidden layers being connected with all the neurons of the previous and of the next layers



It is noteworthy that the first historical Rosenblatt's monolayered perceptron implemented a linear classifier, whereas later multilayered perceptrons functioned with nonlinearly separable data.

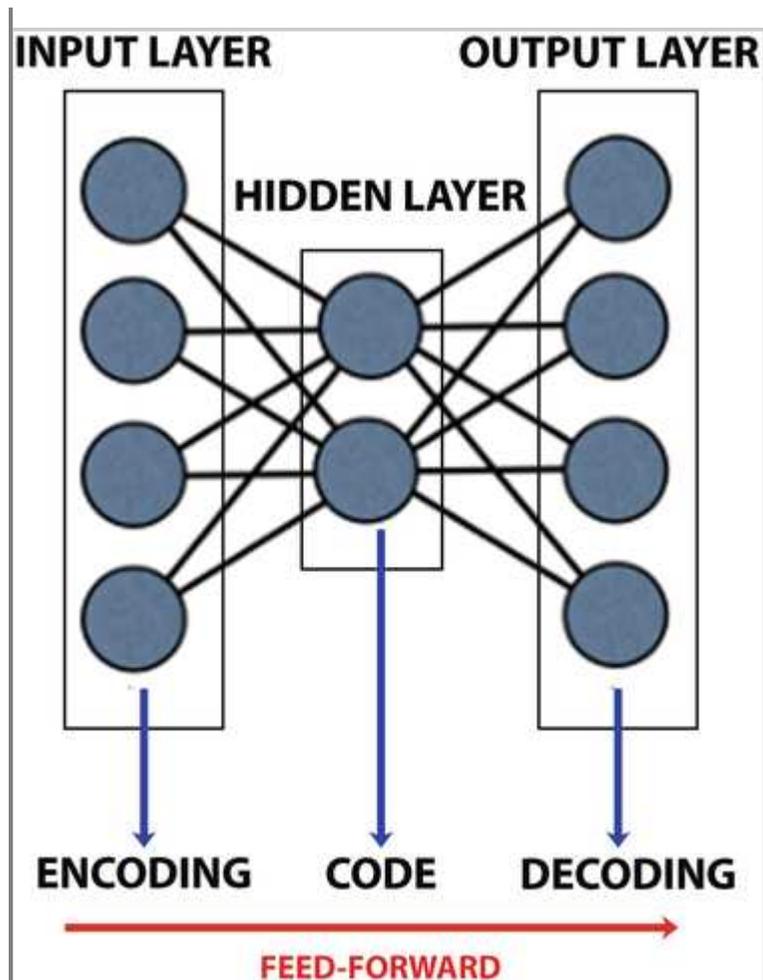
[SVM] can be regarded as a linear classifier partitioning a set of training examples (\mathbf{x}, \mathbf{y}) , or vector, into several classes \mathbf{y} well-delineated by a hyperplane [56, 57]. Examples \mathbf{x} represent the most relevant features of the raw data retained by an instructor for the class attribution. If we restrict our example to two classes \mathbf{y} denoted by -1 and $+1$, the equation of the boundary between these two classes, equivalent to a line, is given by: $\mathbf{f}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{b} = 0$, where \mathbf{w} is the discriminant vector normal to that line, and the function of decision assigning the data \mathbf{x} to its corresponding class \mathbf{y} is: $\mathbf{y} = \text{sign}(\mathbf{w}^T \mathbf{x} + \mathbf{b}) \in \{-1; +1\}$. Moreover, defining optimally the function: $\mathbf{y} = \mathbf{f}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{b}$, which requires to minimize the classification error and the model complexity, supposes to accurately adjust the parameters $(\mathbf{w}^T, \mathbf{b})$. This optimization step relies on the maximum-margin approach. The boundary separating the two classes is the line bisecting the margin demarcated by the closest points belonging to the opposite classes, called [support vectors], characterized by: $\{\mathbf{x} \mid \mathbf{f}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{b} = -1\}$ and $\{\mathbf{x} \mid \mathbf{f}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{b} = +1\}$ (Fig. 2.7). Therefore, the optimizing procedure only takes into account the support vectors and seeks maximization of the geometric margin, which is equivalent to minimizing $\|\mathbf{w}\|^2$ under the following constraint: $\mathbf{y} \cdot (\mathbf{w}^T \mathbf{x} + \mathbf{b}) \geq 1$. Finally, when data are not linearly separable, some mathematical manipulations, known as the "kernel trick," enable to convert easily a nonlinear boundary into a hyperplane by transforming nonlinearly the data space into a higher dimensional feature space, without explicitly calculating the nonlinear mapping. In conclusion, [SVM] can be used for classification of linearly and nonlinearly separable data. Since SVM computation is based on convex optimization, it is guaranteed to converge to global minimum. Thus, SVM appears particularly robust against overtraining when few samples of high dimensional data are processed. [SVM] was successfully tested for disease identification and cerebral state classification [58]. Davatzikos et al. [59] demonstrated the ability of a nonlinear SVM to distinguish

of the last hidden layer with the desired output within a loss layer used during the training phase.

A trainable autoencoder (AE) corresponds to a three-layer, feed-forward neural network mapping, deterministically or stochastically, an input to an output with a minimal amount of distortion. In other words, its hidden layer learns how to encode a compressed representation of the inputs transmitted by the (visible) input layer, by minimizing the quadratic error, and this latent representation, or code, allows the most reliable reconstruction of the inputs by the decoder layer (Fig. 2.8). Broadly speaking, AE accomplish (unperfect) copying tasks. Importantly, AE of lower dimension than the input (and output) layer permits the hidden layer to discriminate the most salient characteristics of the input data. AE can, therefore, be used for feature detection and dimension reduction, and sparse AE behave as classifiers. The computational ability of such ML can be augmented by hierarchically stacking AE so that the output of the previous AE serves as input of the next AE. In case of such SAE, the training procedure is based on a greedy-layer-wise unsupervised pretraining method that requires to train one layer at a time, the input of the next layer being the output of the previous one, using a backpropagating algorithm. A final supervised learning step can be added for improvement of the whole circuit performance.

Fig. 2.8

Autoencoder (AE). A (shallow) AE is a three-layer feed-forward network encoding the input data in the hidden layer, of lower dimension, and decoding the corresponding latent representation in the output layer

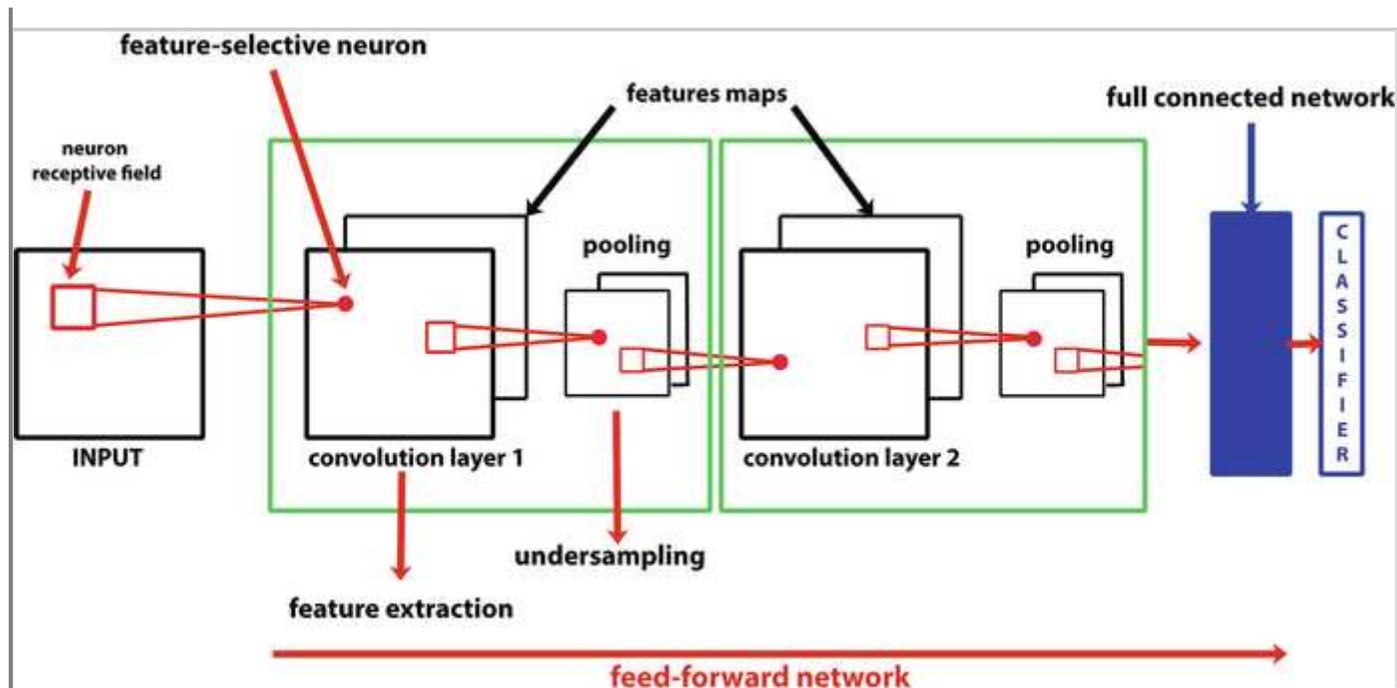


CNN constitutes one of the most influential ML devoted to 2-D and 3-D image recognition, and more generally, to “grid-like topological” data. It forms a subclass of feed-forward multilayered neural network (perceptron). Its architecture includes two types of hidden layers. First, convolution layers act as simple feature detectors (first hidden layer) and more complex feature combination detectors (higher layers) applied to the output of the previous layer (Fig. 2.9). For instance, if the input is an image (voxels), three successive convolutional layers would detect, respectively: (1) simple oriented edges; (2) apparent contours such as lines, curves, or corners; and (3) parts of object. Second, contrary to perceptron, a neuron of **CNN** is not connected to all the neurons of the previous layer. Each neuron of a convolutional layer analyzes a local spatial region of the image, so that a neuron is linked to a local subset of neurons, defining its receptor field. The neurons sensitive to the same feature share the same weight w and realize a spatial paving of the whole incoming data

volume. They, thus, implement a feature-specific filtering (convolving) with a learnable kernel, sliding over the image data and generating feature-specific activation maps after nonlinear processing by an activation function, such as rectified linear unit. Second, pooling layers periodically interspaced and connected with convolution layers yield to subsampling, to “merge semantically similar features into one” [55] and to subsequent translational invariance. This processing contributes to reduce significantly the computational cost and to avoid overtraining. With such architecture, CNN is a powerful tool for object categorization in an image but needs a large sample of training data. In neuroimaging, Brosch et al. [61] successfully segmented lesions of multiple sclerosis using CNN. And other study pointed out the ability of CNN to distinguish resting-state fMRI between normal, mild cognitive impairment patients and Alzheimer disease patients [62], or normal and schizophrenia patients [63].

Fig. 2.9

Schematic representation of a convolutional neural network (CNN) implementing a representation learning algorithm. CNN is composed of the concatenation of several feed-forwardly connected blocks. Each block comprises a convolution layer and pooling layers, which perform feature extraction and undersampling (for dimensional reduction and more invariant representation), respectively. Each “deeper” block processes more and more compressed, complex and abstract conjunction of features. The last block of the chain is represented by a full-connected network and an output layer carrying out data classification. Note that a neuron in a convolution layer acts as a specific local feature detector within its receptive field. Mathematically speaking, these neurons perform first a convolution and afterward a nonlinear processing, such as linear rectifying, with respect to their inputs. All the neurons sharing the same feature (and the same weights) and whose receptive fields, all together, cover the whole previous input layer generate activation feature-specific maps. Therefore, each convolutional layer encompasses several feature maps. CNN are particularly well designed for image analysis and object and pattern recognition



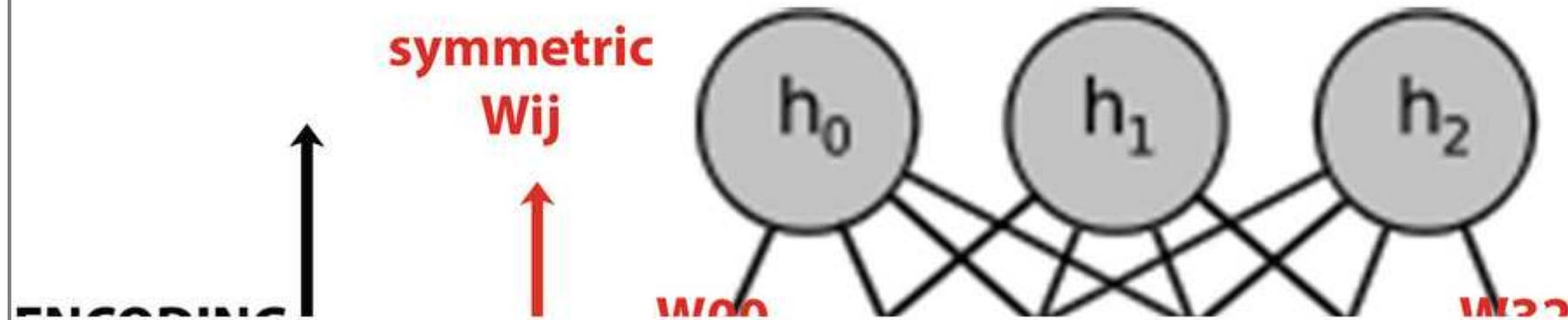
Other **deep-learning algorithms** are based on variant architectures such as DBN, a stochastic neural network composed of stacked restricted Boltzmann machines (**RBM**), using a contrastive divergence algorithm during the training, or such as RNN capable of processing sequential (temporal) data (language, speech, or writing). RBM is a two-layer network composed of a visible (input) layer and one hidden layer linked by symmetric connectivities, but neurons belonging to the same layer are not interconnected (Fig. 2.10). Thus, stochastic binary neurons of the hidden layer elaborate a latent probabilistic representation of the data, and, afterward, this later representation is fed backward in the visible layer to construct a generative model of the data. **RBM** behaves as a stochastic AE. The training algorithm aimed at reducing the difference between the probability distribution of the reconstructed and the original data vectors, by iteratively adjusting the connection weights. Recently, Plis et al. [64] demonstrated that (shallow) DBM and ICA equivalently identified resting-state intrinsic brain networks. Suk et al. [65] managed to discriminate between normal, mild cognitive impairment converter and nonconverter patients and Alzheimer patients, applying DBM processing to MRI, PET, and biochemical data. It could be also assumed that RNN could capture dynamical recruitment and synchronization of neural networks across time.

Fig. 2.10

Restricted Boltzmann machine (**RBM**). Shallow RBM is a two-layer neural network composed of stochastic neurons distributed in an input layer and a hidden layer and linked by weighted symmetric connections. However, there are no intralayer connections between neurons. A latent representation of the data is computed in the hidden layer, which finally is reconstructed, in a second time, in the input layer. RBM implements a probabilistic energy-

based learning algorithm

HIDDEN LAYER



Finally, the choice of the ML depends on the specific characteristics of the data to be studied and on the results of the ML calculation: data categorization or statistical model. Therefore, after the training phase, it appears necessary to evaluate the performance of the algorithm in terms of accuracy, sensitivity, ability to properly generalize, reproducibility, robustness to noise, and adequate adjustment of hyper-parameters (i.e., parameters of the network tuned by the

instructor), for instance, and, among different tested algorithms, to select the algorithm generating the most realistic output. Cross-validation procedures can be exploited for performance estimation and algorithm selection [57].

ML in Neuroimaging In a recent review of application of ML to structural and functional neuroimaging (MRI and PET data), Viera et al. [54] have concluded that ML provides a promising method to precisely classify some neurologic and psychiatric troubles, such as Alzheimer disease, attention-deficit/hyperactivity disorders, schizophrenia, spinocerebellar ataxia, and temporal epilepsy, as well as pain [66], and to predict disease state and disease evolution. They emphasized that deep learning seemed to exhibit better performance than SVM. However, further studies with larger cohorts are required, especially for deep learning. Finally, different ML algorithms can be combined to increase their performance and can be applied to multimodal data (images, biochemical data, clinical data, for instance). For fMRI and, especially, rsfMRI, ML proves to be efficient to localize and to identify activated networks, as well as to differentiate resting-state networks and subtle nonlinear relationships between specific brain areas. However, a clear methodology remains to be established for secure and efficient application of ML to fMRI [67].

Conclusion and Perspectives in Clinical Neuroimaging

In summary, clinical brain imaging benefits from advanced imaging and post-processing methods developed in fundamental neuroscience research. Voxel-based morphometry and measurement of cortical thickness enable to identify and to follow up cortical neuroplasticity and atrophy or disease progression, while tractography deals with white matter and, especially, anatomical tracts. FMRI and rsfMRI detect functional impairments or reconfiguration in specific task-recruited and canonical cerebral networks and can also track time-varying changes in functional brain connectivity using time-window analysis, such as windowed Fourier or wavelet transforms. Multiband acquisition allows for accelerating signal recording with echo time <1 s, improving time resolution of fMRI. Advanced mathematical post-processings complement these static or dynamical anatomo-functional data by objectivizing and quantifying disease-related modulatory, causal, and topological alterations. Machine learning and especially representation learning algorithms used in deep learning, applicable to structural MRI, fMRI, and rsfMRI, will also allow automatic and accurate identification of subtle normal or impaired cerebral patterns.

Moreover, utilization of ultrahigh field (>7 T) would not only increase significantly spatial resolution (μm) for morphological images but also would discriminate activation of cortical columns. This high-resolution fMRI would also permit to differentiate feed-forward thalamic activation of layer IV of cortical column from feed-back afferents impinging on supra- and infragranular layers [68]. This so-called **laminar fMRI** will open in vivo a new window, at the mesoscopic level, to computation of the functional unit of the neocortex and to the top-down and bottom-up influences upon this intracolumnar information processing.

References

1. Raichle ME (1998) Behind the scenes of functional brain imaging: a historical and physiological perspective. *Proc Natl Acad Sci U S A* 93:765–772
2. Hillman EMC (2014) Coupling mechanism and significance of the BOLD signal: a status report. *Annu Rev Neurosci* 37:161–181

3. Buxton RB (2013) The physics of functional magnetic resonance imaging (fMRI). *Rep Prog Phys* 76(9):096601
4. Drake CT, Iadecola C (2007) The role of neuronal signaling in controlling cerebral blood flow. *Brain Lang* 102:141–152
5. Tallini YN, Brekke JF, Shui B, Doran R, Hwang S-M, Nakai J, Salama G, Segal SS, Kotlikoff MI (2007) Propagated endothelial Ca^{2+} waves and arteriolar dilation in vivo. *Circ Res* 101:1300–1309
6. Marelli SP (2001) Mechanisms of endothelial P2Y_1 - and P2Y_2 -mediated vasodilation involve differential $[\text{Ca}^{2+}]_i$ responses. *Am J Physiol Heart Circ Physiol* 28:H1759–H1766
7. Fergus A, Lee KS (1997) GABAergic regulation of cerebral microvascular tone in the rat. *J Cereb Blood Flow Metab* 17:992–1003
8. Li J, Iadecola C (1994) Nitric oxide and adenosine mediate vasodilatation during functional activation in cerebellar cortex. *Neuropharmacology* 33:1453–1461
9. Magistretti PJ, Pellerin L (1996) Cellular bases of brain metabolism and their relevance to functional brain imaging: evidence for a prominent role of astrocytes. *Cereb Cortex* 6:50–61
10. Magistretti PJ, Pellerin L (1999) Cellular bases of brain metabolism and their relevance to functional brain imaging: evidence for a prominent role of astrocytes. *Philos Trans R Soc Lond Ser B Biol Sci* 354:1155–1163
11. Ogawa S, Menon RS, Tank DW, Kim SG, Merkle H, Ellermann H, Ugurbil K (1993) Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophys J* 64(3):803–812
12. Buxton RB, Griffeth VE, Simon AB, Moradi F (2014) Variability of the coupling of the blood flow and oxygen metabolism responses in the brain: a problem for interpreting BOLD studies but potentially a new window on the underlying neural activity. *Front Neurosci* 8:139
13. Davis TL, Kwong KK, Weiskoff RM, Rosen BR (1998) Calibrated functional MRI: mapping the dynamics of oxidative metabolism. *Proc Natl Acad Sci U S A* 95(4):1834–1839
14. Kim SG, Ogawa S (2012) Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. *J Cereb Blood Flow Metab* 32:1188–1206
15. Shih Y-YI, Chen C-CV, Lin Z-J, Chiang Y-C, Jaw F-S, Chen Y-Y, Chang C (2009) A new scenario for negative functional magnetic resonance imaging

- signals: endogenous neurotransmission. *J Neurosci* 29(10):3036–3044
16. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157
 17. Wang J, Aguirre GK, Kimberg DY, Roc AC, Li L, Detre JA (2003) Arterial spin labeling perfusion fMRI with very low task frequency. *Magn Reson Med* 49(5):796–802
 18. Chen JJ, Jann K, Wang DJJ (2015) Characterizing resting-state brain function using arterial spin labeling. *Brain Connect* 5(9):527–542
 19. Buxton RB (2016) Beyond BOLD correlations: a more quantitative approach for investigating brain networks. *J Cereb Blood Flow Metab* 36(3):461–462
 20. Dai W, varma G, Scheidegger R, Alsop DC (2016) Quantifying fluctuations of resting state networks using arterial spin labeling perfusion MRI. *J Cereb Blood Flow Metab* 36(3):463–473
 21. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34(4):537–541
 22. Gusnard DA, Raichle MES (2011) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694
 23. Raichle EM (2015) The restless brain: how intrinsic activity organizes brain function. *Philos Trans B* 370:1–11
 24. Fox MD, Raichle M (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711
 25. Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. *Proc Natl Acad Sci U S A* 104:13170–13175
 26. Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, Barnes GR, Smith SM, Moris PG (2011) Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proc Natl Acad Sci U S A* 108(40):16783–16788
 27. Shmuel A, Leopold DA (2008) Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: implications for functional connectivity at rest. *Hum Brain Mapp* 29(7):751–761

28. Leopold DA, Maier A (2012) Ongoing physiological processes in the cerebral cortex. *NeuroImage* 62:2190–2200
29. Greco G, Corbetta M (2011) The dynamical balance of the brain at rest. *Neuroscientist* 17:107–123
30. Margulies DS, Böttger J, Long X, Lv Y, Kelly C, Schäfer A, Goldhahn D, Abbushi A, Milham MP, Lohmann G, Villringer A (2010) Resting developments: a review of fMRI post-processing methodologies for spontaneous brain activity. *MAGMA* 23(5–6):289–307
31. Beckmann CF (2012) Modelling with independent components. *NeuroImage* 62:891–901
32. Calhoun VD, Liu J, Adali T (2009) A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic and ERP data. *NeuroImage* 45:S163–S172
33. Zou Q-H, Zhu C-Z, Yang Y, Zuo X-N, Long X-Y, Cao Q-J, Wang Y-F, Zang Y-F (2008) An improved approach to detection of amplitudes of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J Neurosci Methods* 172(1):137–141
34. Zang Y, Jiang T, Lu Y, He Y, Tian L (2003) Regional homogeneity approach to fMRI data analysis. *NeuroImage* 22(1):394–400
35. Kelly RE, Alexopoulos GS, Wang Z, Gunning FM, Murphy CF, Morimoto SS et al (2010) Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. *J Neurosci Methods* 189:233–245
36. Kalcher K, Huf W, Boubela RN, Filzmoser P, Pezawas L, Biswal B et al (2012) Fully exploratory network independent component analysis of the 1000 functional connectomes database. *Front Hum Neurosci* 6:1–11
37. Habas C, Kamdar N, Nguyen D, prater K, Beckmann CF, Menon V, Greicius MD (2009) Distinct cerebellar contribution to intrinsic connectivity networks. *J Neurosci* 29(26):8586–8594
38. Chen AC, Oathes DJ, Chang C, Bradley T, Zhou Z-W, Williams LM et al (2013) Causal interactions between fronto-parietal central executive and default-mode networks in humans. *Proc Natl Acad Sci U S A* 110(49):19944–19949
39. Fox MF, Corbetta M, Snyder AZ, Vincent J, Raichle M (2006) Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc Natl Acad Sci U S A* 103:10046–10051
40. Hoff GEA-J, Van de Heuvel MP, Benders MJNL, Kersbergen KJ, de Vries LSD (2013) On the development of functional brain connectivity in the young brain. *Front Hum Neurosci* 7:650

41. Antonenko D, Flöel A (2014) Healthy aging by staying selectivity connected: a mini-review. *Gerontology* 60:3–9
42. Kelly C, Castellanos FX (2014) Strengthening connections: functional connectivity and brain plasticity. *Neuropsychol Rev* 24:63–76
43. Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD (2014) Tracking whole-brain connectivity dynamics in the resting state. *Cereb Cortex* 24(3):663–676
44. Fox MD, Greicius M (2010) Clinical applications of resting state functional connectivity. *Front Syst Neurosci* 4:1–13
45. Rosazza C, Minati L (2011) Resting-state brain networks: literature review and clinical applications. *Neurol Sci* 32:773–785
46. Friston KJ (2011) Functional and effective connectivity: a review. *Brain Connect* 1(1):13–36
47. Penny WD, Stephan KE, Mechelli A, Friston KJ (2004) Modelling functional integration: a comparison of structural equation and dynamic causal models. *NeuroImage* 23(Suppl 1):S264–S274
48. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *NeuroImage* 19(4):1273–1302
49. Ed B, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10:186–198
50. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52:1059–1069
51. Guye M, Bettus G, Bartolomei F, Cozzone PJ (2010) Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *MAGMA* 23:409–421
52. Redcay E, Moran JM, Mavros PL, Tager-Flusberg H, Gabrieli JDE, Whitfield-Gabrieli S (2013) Intrinsic functional network organization in high-functioning adolescents with autism spectrum disorder. *Front Hum Neurosci* 7(573):1–11
53. Wang S, Summers RM (2012) Machine learning and radiology. *Med Image Anal* 16(5):933–951
54. Vieira S, Pinaya WHL, Mechelli A (2017) Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci Behav Rev* 74:58–75
55. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444

56. Wernick MN, Yang YY, Braaankov JG, Yourganov G, Strother SC (2010) Machine learning in medical imaging. *IEEE Signal Process Mag* 27(4):25–38
57. Lemm S, Blankertz B, Dickhaus T, Müller K-R (2011) Introduction to machine learning for brain imaging. *NeuroImage* 56:387–399
58. Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* 19(2):261–270
59. Davatzikos C, Ruparel K, Fan Y, Shen D, Acharyya M (2005) Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *NeuroImage* 28(8):663–668
60. Shen D, Wu G, Suk H-I (2017) Deep learning in medical images analysis. *Annu Rev Biomed Eng* 19:221–248
61. Brosch T, Tam R, Alzheimer’s Disease Neuroimaging Initiative (2013) Manifold learning of brain MRIs by deep learning. In: *International conference on medical image computing and computer-assisted intervention*, Springer Berlin Heidelberg, pp 633–640
62. Suk H-I, Lee S-W, Shen D (2014) Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* 101:569–582
63. Kim J, Calhoun VD, Shim E, Lee J-H (2016) Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *NeuroImage* 124(Pt A):127–146
64. Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt H, Long JD, Johnson HJ, Paulsen JS, Turner JA, Calhoun VD (2014) Deep learning for neuroimaging: a validation study. *Front Neurosci* 8:229
65. Suk H-I, Lee S-H, Shen D (2015) Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Struct Funct* 220(2):841–859
66. Rosa MJ, Seymour B (2014) Decoding the matrix: benefits and limitations of applying machine learning algorithms to pain neuroimaging. *Pain* 155:864–867
67. Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage* 15(1 suppl):S199–S209
68. Self MW, van Kerkoerle T, Goebel R, Roelfsema PR (2017) Benchmarking laminar fMRI: neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. *Neuroimage*. 2017 Jun 23. pii: S1053-8119(17)30517-7. doi: 10.1016/j.neuroimage.2017.06.045.

[Epub ahead of print]

View publication stats