Published in final edited form as:

Nat Neurosci. 2017 February 23; 20(3): 304–313. doi:10.1038/nn.4499.

Computational approaches to fMRI analysis

Jonathan D Cohen[1,2], **Nathaniel Daw**[1,2], **Barbara Engelhardt**[3], **Uri Hasson**[1,2], **Kai Li**[3], **Yael Niv**[1,2], **Kenneth A Norman**[1,2], **Jonathan Pillow**[1,2], **Peter J Ramadge**[4], **Nicholas B Turk-Browne**[1,2], and **Theodore L Willke**[5]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, USA

[2]Department of Psychology, Princeton University, Princeton, New Jersey, USA

[3]Department of Computer Science, Princeton University, Princeton, New Jersey, USA

[4]Department of Electrical Engineering, Princeton University, Princeton, New Jersey, USA

[5]Intel Labs, Intel Corporation, Santa Clara, California, USA

Abstract

Analysis methods in cognitive neuroscience have not always matched the richness of fMRI data. Early methods focused on estimating neural activity within individual voxels or regions, averaged over trials or blocks and modeled separately in each participant. This approach mostly neglected the distributed nature of neural representations over voxels, the continuous dynamics of neural activity during tasks, the statistical benefits of performing joint inference over multiple participants and the value of using predictive models to constrain analysis. Several recent exploratory and theory-driven methods have begun to pursue these opportunities. These methods highlight the importance of computational techniques in fMRI analysis, especially machine learning, algorithmic optimization and parallel computing. Adoption of these techniques is enabling a new generation of experiments and analyses that could transform our understanding of some of the most complex—and distinctly human—signals in the brain: acts of cognition such as thoughts, intentions and memories.

Techniques for human brain imaging grew out of radiology, initially involving radioactive tracers (for example, positron-emission tomography (PET)), followed by the momentous discovery that MRI could measure intrinsic hemodynamic signals linked to neural activity (functional MRI or fMRI)[1–3]. Initial analyses of radiotracer-based data quantified absolute activity, taking into account the pharmacodynamics of tracer delivery and clearance, and thus they were complex and specialized. The application of subtractive methods—measuring relative activity for experimental versus control conditions—made analysis more

Correspondence should be addressed to N.B.T.-B. (ntb@princeton.edu).

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

All authors helped conceive the manuscript and wrote one or more sections. N.B.T.-B. edited the manuscript. N.B.T.-B. revised the manuscript with input from J.D.C., K.A.N. and P.J.R. Author order was determined alphabetically.

straightforward[4]. The earliest methods compared the measurements at each location (volumetric pixels or 'voxels') statistically using *t*-tests. This estimated the change in each voxel's activity in response to the experimental manipulation and was used to construct a 'map' of such statistics, indicating the distribution of activity over the brain[5].

This simple approach faced limitations. First, it involved binary comparisons, which can be underpowered when identifying continuous processes; this was partly addressed by parametric designs, which used regression to identify voxels responsive in a predicted way[6]. Second, early approaches were compromised by the delayed and protracted course of the hemodynamic response relative to neural activity; deconvolution methods were developed to account for this[7], often by assuming a consistent function across regions and individuals. Third, the scale of brain imaging data and the number of statistical comparisons created a high risk of false discoveries (Type I errors); in response, methods were developed that exploit priors about the data (for example, spatial contiguity) when correcting for multiple comparisons[8,9].

These developments, together with consensus about the most important kinds of preprocessing (motion correction, slice-time correction, filtering in space and time, and anatomical alignment across individuals), led to the creation of standard software toolboxes that have been in widespread use for two decades[10–13]. They have had a dramatic and positive impact on fMRI research, standardizing practices in the field and facilitating the dissemination of methods. With the advent of novel mathematical, statistical and computational techniques in science at large, and with the dramatic growth of technology and computing power, these tools have continued to evolve and new tools have emerged. In what follows, we discuss some of these newer developments, with a focus on the increasing importance of computation.

Selective review of advanced fMRI analyses

The central role of computation in neuroscience is evident in the approaches to fMRI analysis developed over the last decade. These approaches built on techniques and concepts from computer science and engineering (for example, machine learning, graph theory, control theory), as well as from advances in these fields (for example, software and hardware optimization and parallelization) that allow these and more traditional approaches to run more efficiently and at larger scale. This has had both quantitative and qualitative impacts on the science that is possible. Here we review three modern analysis approaches that have been computationally informed in this way.

Multivariate analysis

As opposed to univariate methods, which examine individual voxels or regions, multivoxel pattern analysis (MVPA) considers spatial patterns of activity over ensembles of voxels to recover what information they represent collectively[14,15]. These methods are effective, although the reasons for this are still debated. On one hand, MVPA may be sensitive to information at a subvoxel scale, insofar as neuronal populations are distributed heterogeneously over voxels such that there is subtle variation in the tuning of individual voxels[16]. On the other hand, the spatial distribution of information may reflect a larger-scale

map being sampled by voxels, in which case MVPA would not provide finer-grained evidence of neural selectivity[17]. A valuable way forward on this issue is to model how neuronal activity manifests at the level of voxels, including examining how voxel size, the distribution of cognitive functions and the vasodilatory response impact what information is retained in simulated voxel activity patterns[18].

The most common form of MVPA uses classifiers from machine learning (Fig. 1a), typically linear models such as logistic regression. The classifier learns a weight for each voxel, and these weights together determine the decision boundary between experimental conditions. During training, weights are adjusted to maximize how well the boundary separates the conditions. To avoid overfitting noise, the complexity of the classifier is often constrained, including with regularization, which punishes undesirable or unlikely solutions (for example, large weights). When tested on new data, the weights are used to calculate a weighted sum of voxel activity, which is compared against the boundary to guess the class. Classifiers can be applied over the whole brain[19], in spatial moving windows ('searchlights')[20] or in regions of interest (ROIs).

Classifier-based MVPA has been leveraged to derive trial-by-trial measures of internal cognitive states, such as what a participant is attending to, thinking about or remembering. Such studies often track states that are easy to discriminate in fMRI (for example, face and scene processing)[21–24]. However, some studies have found success decoding more fine-grained states, such as which specific item is being predicted[25], replayed[26] or recollected[27]. One challenge for this general approach is that classifiers are opportunistic and exploit any discriminative variance, making it especially important to control for factors that are confounded with the variables of interest, such as task difficulty and reaction time[28].

The second major type of MVPA focuses on the similarity of voxel patterns (Fig. 1b). Activity patterns are viewed as points in a high-dimensional voxel space, where the distance between points indicates their similarity. Rather than dividing the space with a classifier, it is summarized as a matrix of pairwise distances[29]. The structure of the matrix can reveal what information is encoded in a region by comparing it to other similarity matrices, such as from human judgments or computational models[30–32]. Similarity-based MVPA has also been used to track how learning influences neural patterns[33–39]. One caveat for this general method is that all voxels are weighted equally, unlike in classifier-based MVPA, and thus there is a risk of contamination from uninformative or noisy features. Another caveat is that pattern similarity can be easily confounded, including by univariate activity[40] and temporal proximity[41]: in such cases, effects on similarity might be interpreted as neural patterns converging or diverging in representational space, when in fact the underlying structure of the neural patterns has not changed.

Real-time analysis

In a normal workflow, fMRI data are collected, transferred from the scanner to a server and analyzed offline over weeks, months or years. What could be gained from performing analysis during rather than after data collection, obtaining the results in seconds? This question has driven interest in real-time fMRI[42,43]. As a research tool, real-time fMRI has opened up intriguing opportunities for training and/or novel experimental designs. In

particular, by analyzing data on the fly, the results can be used to adjust the ongoing experiment (Fig. 2).

The most widespread kind of adjustment involves trying to influence the participant via feedback about their brain activity ('neurofeed-back'). Inspired by the tradition of using EEG for biofeedback, the goal is often to train a participant to increase or decrease activity in a region of the brain underlying some cognitive process or disorder. fMRI neurofeedback has been used clinically, such as for chronic pain[44], depression[45] and addiction[46], as well as to understand basic cognitive functions, such as perceptual learning[47] and sustained attention[48].

It is notable that real-time fMRI and neurofeedback have been around since the early days[49–51], and although there has been some success in the interim[44,52], a major resurgence is underway. One possible reason is that the field has a better understanding of when neurofeedback works, with recent studies of which mental strategies are effective[53] and which brain regions are more controllable[54] versus involved in controlling[55]. The parallel growth of MVPA may have also contributed, as it might be easier to control specific contents of mind (reflected in distributed representations) than the mean activity level of regions that are often linked to multiple cognitive functions[42,47]. Moreover, incorporating feedback into cognitive tasks in a closed-loop manner (for example, via stimulus contrast or task difficulty) may feel more natural to participants and enable a broader range of real-time designs than the scales or gauges used typically[48]. Nevertheless, fMRI neurofeedback will always be limited by the hemodynamic lag, and thus feedback may be most informative about cognitive and neural processes that drift slowly (for example, attention, motivation and learning) and are therefore likely to be in a similar state at the time of feedback, despite the delay.

In neurofeedback studies, real-time results are used by the participant to change their strategy or behavior. The other major classes of real-time fMRI place more emphasis on what the experimenter does with the results. In 'triggering' designs, the level of activity in a brain region is monitored by the experimental control apparatus, and trials are initiated when activity is low or high, with predictions of different behavioral consequences in these two scenarios[56]. This may enhance causal inference in fMRI, as brain activity serves as an independent variable, potentially making it possible to understand (with appropriate control regions) whether a given region is sufficient for the behavior. In 'adaptive' designs, the experimenter determines not whether to present a trial—trials occur at regular intervals irrespective of brain activity—but rather the content of the next trial (for example, the stimulus or task). This has been done with the goal of characterizing the tuning properties of the visual system, by adjusting stimulus parameters until an area's response is maximal[57], but could also be used to examine a wide variety of systems (for example, those involved in attention or memory).

Model-based analysis

A key use of computational models in fMRI is to define hypothetical signals of interest. Processes close to the periphery, such as visual perception, often involve concrete, quantifiable variables that are relatively straightforward to conceptualize, manipulate and

measure. For example, it seems clearer how to design an experiment to test a brain region's involvement in color vision than to test for more abstract constructs such as prediction error or confidence. The advantage is not just that color is more tangible but also that the psychology and neuroscience of vision have long been guided by a firm computational foundation, via formalisms such as signal detection theory. Such theories, originally from engineering but now deeply ingrained in experimental design and analysis, specify key steps in perception and how their operation can be assessed.

In contrast, higher-level aspects of cognition, such as decision, valuation, control and social interactions, were relatively slower to benefit from computational theories that suggest which quantities to manipulate and measure. Models of these processes (for example, reinforcement learning, decision theory, Bayesian inference and game theory) have seen increasing use as tools to develop precise hypotheses about the underlying computations. For instance, reinforcement learning specifies how choice outcomes influence future decisions and game theory describes how social agents respond to each other's actions. These hypotheses can in turn be used to generate predictions about neural signals (Fig. 3). Specifically, if a model correctly specifies a computation in the brain, that model can be used to estimate time-varying signals for otherwise subjective, hidden variables (for example, value expectation, attention signals, accumulated evidence), the correlates of which can be sought in the brain[58].

Model-based fMRI allows researchers to go beyond localizing model variables in the brain. Once their location is known, these signals can be read out in later experiments and used to estimate parameters independently of behavior (for example, loss aversion[59]), to arbitrate which computational processes are being used by a participant via model comparison (for example, in anticipating opponents' behavior[60], computing decision variables[61] or allocating dimensional attention[62]) and for the kinds of real-time designs discussed above. Recent work has pushed the boundary further, particularly by combining model-derived time series with other modes of fMRI analysis, including visual category decoding[63] and repetition suppression[64].

Approaches for scaling up advanced fMRI analyses

The amount of empirical data and the complexity of theoretical models are both continuing to increase at a rapid pace. A central challenge for neuroscience is to develop methods that can scale gracefully with this growth. Here we discuss scalable approaches that seek to address this problem, as well as technological advances that could help facilitate solutions.

Finding the signal: methods for identifying meaningful variance

fMRI analysis can be hamstrung by the small number of observations per participant relative to the complexity of the data. If every voxel is considered a dimension of variation, then activity patterns over voxels can be described as points in this high-dimensional space. Given that the volume of this space expands dramatically with the number of dimensions (voxels), a smaller number of observations (volumes) than voxels means that the points corresponding to the activity patterns for these observations will fill the space very sparsely. This makes statistical analysis difficult and underpowered, for example leading to poorly

placed decision boundaries in MVPA, which in turn impairs classification performance. This 'curse of dimensionality' is only exacerbated by the fact that blood oxygenation level-dependent (BOLD) activity is very noisy. Given these challenges, the models fitted to fMRI data need constraints to help them to find the 'needles' of signal embedded in the much larger haystack of noise. Here we describe techniques that, in concert, may improve our ability to identify meaningful cognitive signals.

One such technique is shared response modeling (SRM[65]; see also hyperalignment[66]), which projects fMRI responses from each participant into a low-dimensional space that captures temporal variance shared across participants (Box 1; Fig. 4–5). If participants are given the same stimulus or task sequence (for example, a movie), which leads their brains through a series of cognitive states (for example, visual, auditory, semantic), then identifying shared variance has the effect of highlighting variance related to these states. An added benefit is that SRM helps address the data starvation problem above: because the SRM space is by definition shared across individuals, data from multiple participants can be combined prior to MVPA or other analyses. Cross-participant decoding is possible without SRM[67–70] but may be limited to cognitive states whose neural representations are coarse and thus tolerant of misalignment. Indeed, SRM improves MVPA precisely by aligning fine-grained spatial patterns within local regions[71]. Moreover, beyond improving alignment and increasing the sensitivity of other analyses, the output of SRM itself can be informative. For example, it has been used to estimate the dimensionality with which the posterior medial cortex represents movies during perception and memory recall[72].

Box 1

Shared response model (SRM)

After standard alignment, fMRI data can be aggregated at the group level by averaging the values at each voxel across participants. Although this reduces inter-subject noise, variation in the anatomical locations of functional signals across participants blurs estimates of their shared responses. SRM offers an alternative approach by jointly factoring each participant's data into a shared set of feature time series and subject-specific topographies for each feature (Fig. 4).

The simplest use of SRM is for extracting a shared response across an anatomical ROI. Used this way, SRM and related methods can yield significant gains in sensitivity for group-level inference[65,66]. For example, which short segment of a movie is being watched can be classified with many times greater accuracy from fMRI data after functional versus anatomical alignment[65,71]. Moreover, text annotations of movie segments based on fMRI are consistently better, across ROIs and analysis parameters, after SRM[98]. Applying SRM to a large swath of the brain means that all voxels within the region contribute to the final derived metric. This can conflict with the goal of associating spatially local activity with specific cognitive functions. To address such issues, SRM can be applied in small overlapping searchlights to obtain localized metrics of shared information[71,99].

SRM is computed using a subset of the available fMRI data, with the number of features, k, determined using cross-validation. Naturalistic stimuli such as movies and stories are

often used to generate such training data, though any study design in which participants perform the same sequence of trials—or for which a common sequence can be spliced together from the same set of trials—could be used (for example, a battery of cognitive tasks). SRM highlights the sources of variance elicited by the stimuli or trials that are shared across participants in the training data. Held-out test data (including from new participants) are then projected into the shared response space for further analysis. Such test data could be of the same type as the training data, for example, allowing for decoding of new movie segments (Fig. 5). Alternatively, the test data could be from controlled laboratory experiments—there is no requirement for a common trial sequence or set, unlike for the training data—with SRM simply replacing standard alignment in the preprocessing pipeline. As a rule of thumb, SRM will improve sensitivity for detecting a cognitive process of interest in the test data if the training stimuli or trials strongly and variably engage that process in a way that is reliable across participants. One limitation when using SRM for preprocessing is that additional data must be collected for training, reducing the amount of data (and potentially statistical power) related to the principal question of the study.

The flip side of focusing on shared responses is to focus on responses that are idiosyncratic to individuals. Although these responses are excluded in SRM, they are not necessarily noise and may in fact be highly reliable within participants. Indeed, SRM can be used to isolate participant-unique responses by examining the residuals after removing shared group responses, or it can be applied hierarchically to the residuals to identify subgroups[65]. More generally, there is a growing trend toward investigating individual differences as another source of meaningful variance in fMRI[73]. Recognizing that signal exists beyond the average or shared response of a group, such studies exploit idiosyncratic but stable responses to account for previously unexplained variance in brain function, behavioral performance and clinical measures[70,74].

A second promising technique is to impose spatial priors on the brain patterns extracted by fMRI, based on knowledge of how neural representations relevant to cognitive function are organized. One simple but powerful idea is that such representations are realized sparsely: that is, only a subset of voxels is modulated by a given process of interest. However, sparsity alone is insufficient, as cognitively relevant patterns also tend to be spatially structured, such that nearby voxels coactivate. Bayesian hierarchical models are particularly effective at implementing such simultaneously sparse and structured priors. These models support flexible specification of the spatial prior and share statistical strength across separate observations with the same latent structure, such as data from multiple participants.

Topographic factor analysis (TFA) is one such Bayesian approach that exploits structured sparsity[75]: fMRI images are redescribed in terms of a small (sparse) number of localized sources that have a pre-specified functional form (structure), such as a radial basis function. Given a set of fMRI images, TFA infers the number, locations and sizes of sources that best describe the images, as well as source weights that specify how active each source is in each image. Because the number of sources is typically substantially smaller than the number of voxels in an fMRI dataset, computations based on TFA sources can be orders of magnitude

more efficient than computations based on voxels. Of course, spatial priors and dimensionality reduction both run the risk of excluding signals of interest and thus need to be used in conjunction with more exploratory methods.

A third way to extract signal is to compute patterns of covariance between voxels or regions. Such 'functional connectivity' can carry information about the interactions between regions not evident in localized activity[76]. This may be especially true for processes such as attention, in which certain brain regions control or influence other regions. Indeed, functional connectivity has helped reveal mechanisms for visual selection[77], markers of sustained attention ability[74] and networks that support cognitive tasks more generally[78,79].

The challenge is that voxel covariance patterns are several orders of magnitude larger than the raw data (~10^{10} voxel pairs), increasing the size of the haystack that must be searched for the signal needle. One effective solution has been to reduce the scale of the problem by parcellating the brain into a smaller set of larger regions or clusters before analysis[78]. However, this requires assumptions about the right functional 'units' of neural processing, and such decisions can impact results[80]. The techniques described above can also help: focusing on variance shared across participants can clarify connectivity results (Box 2; Fig. 6). Finally, it is possible to analyze covariance patterns at full voxel scale, but this requires computational optimization and parallelization. This is illustrated by full correlation matrix analysis (FCMA), which uses advanced algorithms to compute the pairwise correlation of every voxel with every other voxel over multiple time windows and to train a classifier on these correlations for decoding held-out time windows, all while intelligently splitting these computations over threads, cores and nodes on high-performance computing systems[81]. Some drawbacks of FCMA include its need for substantial computing power and the difficulty in interpreting and visualizing the results (classifier weights on voxel pairs).

Box 2

Intersubject functional connectivity

One of the most widespread types of fMRI analysis is the investigation of functional connectivity (FC), the temporal covariance of BOLD activity across distinct brain regions and how it changes as a function of external input and internal goals[76,78,79]. FC is most commonly computed between a single seed ROI and the rest of that person's brain (Fig. 6a), but it can also capture interactions across all possible pairs of voxels or regions[81].

The term 'functional connectivity' implies that the covariance between regions is driven by their direct interaction. However, this is not always the case, as covariance can be indirectly caused by physiological factors (for example, breathing or heartbeat) that jointly influence regions or by the synchronization of regions to the same external stimulus. Physiological confounds can be controlled by comparing covariance across two or more experimental conditions, as long as the physiological changes are uncorrelated with the design. Stimulus confounds can be addressed in various ways, including by regressing out stimulus-evoked responses and examining 'background connectivity' in the residuals, providing a purer measure of intrinsic interactions[77].

Here, we focus on the complementary problem of isolating stimulus-driven covariance between regions. A new variant of FC called intersubject functional correlation (ISFC)[96] achieves this goal by calculating regional covariance across participants (for example, correlating region A in participant 1 with region B in participant 2; Fig. 6b). Given that intrinsic neural responses cannot systematically align across participants' brains, the covariance pattern produced by ISFC during rest (when no stimulus is present) should be low and statistically insignificant. However, when neural responses lock to a stimulus, ISFC can isolate covariance that is shared across participants.

ISFC is particularly effective at filtering out spontaneous neural responses, which contribute strongly to FC, while improving sensitivity to stimulus-locked processes. To see this, consider four conditions that were scanned with fMRI: listening to a 7-min story, listening to versions of the story scrambled at the sentence level and at the word level, and resting with no stimulus. FC computed within each participant and then averaged across a group of 18 participants was stable across the four conditions despite large differences in stimulus properties (Fig. 6c), suggesting that regional covariance was dominated by intrinsic interactions and was relatively insensitive to dynamic stimulus-induced variance.

These results can be contrasted with ISFC results for the same data, which showed strong differences across conditions (Fig. 6d): no ISFC at rest; for scrambled words, ISFC was observed only in auditory cortex and early language areas that process words; for scrambled sentences, ISFC extended to broader language networks, encompassing Wernicke's and Broca's areas; and for the intact story, ISFC encompassed the full default-mode network. Moreover, ISFC revealed reliable changes in the configuration of the covariance patterns over short time windows as the story unfolded[96]. This ability to track changes in regional synchronization during the processing of real-life information opens up new avenues for linking the dynamics of brain networks to stimulus features and human behavior.

As with SRM, ISFC discounts variance that is idiosyncratic to individuals. Thus, it should be used in concert with FC, especially to capture meaningful noise correlations[77] and stable individual differences[70,74].

Lastly, there may be important cognitive state information in fMRI that does not manifest in voxel activity or pairwise correlations between voxels but rather in higher-order network properties, irrespective of their particular spatial coordinates (for example, does a set of voxels form a 'loop' or circuit?). Such relationships can be characterized and quantified with graph theory and topological data analysis, which are beginning to be applied to fMRI data[82,83].

Knowing where to start: models for guiding and constraining analysis

The new methods described thus far seek to identify and amplify the signal in fMRI data by exploiting advances in statistical and computational methods, disciplined by our growing understanding of generic principles about how information is represented (for example, sparsely) and processed (for example, interactively) in the brain. On their own, however,

these methods are largely exploratory. Although this has the advantage of being unbiased, it fails to exploit a deepening understanding of the functions carried out by the brain.

As in all forms of analysis, useful prior information can greatly reduce the search space and improve the likelihood of detecting a signal of interest. Priors can come not only from general principles of neural organization (as above) but also from more specific information about the structure of particular processes at play. Such structure may be imposed by the stimuli: for example, computational analyses of patterns of word co-occurrence from large online databases have been used to constrain analysis of fMRI data acquired while those words are being perceived[84]. Structure can also arise from hypotheses about underlying cognitive mechanisms, particularly when the mechanisms can be cast in quantitative form. However, most efforts at such model-based analysis (as discussed earlier) have used low-dimensional models with simple, usually linear functional forms. Though a valuable start, such models fall short of capturing the high-dimensional, nonlinear nature of processing in the brain.

More sophisticated models that confront such complexity have a long tradition in cognitive psychology[85,86]. With computational advances and larger training sets, this approach has seen a dramatic resurgence over the past few years, most evidently in the rise of ‘deep learning’ models of perception[87]. The neuroscience community is beginning to harness the power of such models by integrating them into the analysis of neural data[88]. One approach is to generate predictions about the similarity structure of patterns of fMRI activity from the similarity structure of simulated neural activity patterns in the model[30]. With this approach, hypotheses can be generated about circuits or stages of processing by mapping different layers of the model to different brain areas[89,90].

Successes thus far have mostly come from fitting deep learning models to the visual system (cf. ref. 91). Whether biologically inspired models of higher-level cognitive processes provide similarly powerful insights about other brain systems remains an open question. One difficulty is that the ‘ground truth’ for such processes is hard to define: perceptual models can be trained on millions of photographs with unambiguous labels about what objects are contained, but there is no similar corpus of memories or thoughts and no accepted vocabulary or basis set for describing their contents. Even if such a corpus existed, the point-spread function of the hemodynamic response in space and time, as well as our still-imperfect neurophysiological understanding of BOLD activity, complicate the translation of simulated activity in model units to predicted fMRI activity patterns. Indeed, the scale and manner with which neuronal populations are sampled and locally averaged in fMRI voxels can have a big impact on the information carried by voxel activity patterns, although repeated random sampling provides some stability[18]. Continued advances in mathematical and computational methods may help address these problems and license the use of more detailed and realistic models of neural function in the analysis of fMRI data.

Getting the job done: scalable computing for efficient analysis

We have made the case that future analyses need to explore data and exploit models at full scale but have not yet addressed the practical issue of whether this is tractable. The methods above can be extremely computationally demanding and data intensive, with the number of

operations and the memory needed for intermediate products scaling exponentially with the size of the dataset, or worse. For example, SRM requires a matrix inversion equal to the total number of voxels pooled across all participants, and FCMA trains a separate classifier of seed-based whole-brain connectivity for every voxel during feature selection. High-resolution and multiband imaging, combined with an increased focus on jointly considering all participants, worsen the problem. To complete the analysis of a single dataset may require years with a standard implementation, even when deployed on a modern compute server. This would, of course, put real-time applications out of reach, as well as hamper the rate of scientific discovery and progress more generally.

These computational bottlenecks have attracted the interest of computer scientists, not only from machine learning but also from algorithms and systems research[92,93]. The complexity of algorithms can be reduced with efficient mathematical transformations and numerical methods. Algorithms can be further optimized by one to two orders of magnitude by taking advantage of modern hardware like multi-core, manycore and GPU boards; using cache and memory hierarchies more intelligently; improving data staging between processing steps; and exploiting instruction-level parallelism such as single instruction, multiple data (SIMD) and vector floating-point units. Analyses can also be parallelized over networks of computers[94], at the levels of both data and models. In data parallelism, each processor keeps a full replica of the model to train on a portion of the data. In model parallelism, each processor receives a portion of the model to train on all data, sharing parameters to aid convergence. Network parallelization must be designed for minimal and efficient communication, such as with the message passing interface (MPI) protocol. Together, these measures will permit neuroscientists to take full advantage of high-performance computing resources, with the potential of a near-linear speedup per machine on the network. However, this requires that neuroscientists gain the computational expertise to implement these approaches and/or develop close collaborations with computational scientists.

In addition to scaling up neuroimaging analyses with computational techniques, such analyses will benefit by being scaled up in a different way, from the small number of laboratories and institutions with the expertise and equipment to create them to the much broader community of researchers who use fMRI. This requires developing code in a way that can be shared and run by others, as well as standardizing file types and a lexicon for annotating experimental details, so that new data can be analyzed using the code[95]. To facilitate replications, meta-analyses, classroom instruction and personnel training, specific instantiations of the code and parameters should be included in publications, along with the corresponding raw data. Finally, even if code and/or data are shared, many of the future analyses discussed here run efficiently only on large clusters, which are available to just a subset of the user base. This could be ameliorated by converting analyses to a software-as-a-service (SaaS) or ‘cloud’ ecosystem (Box 3; Fig. 7), an approach that has revolutionized many fields and industries. Such developments will again require close collaborations between neuroscientists and computational scientists.

Box 3

Real-time cloud software-as-a-service (SaaS)

The computational demands of fMRI analysis are growing as datasets get larger and algorithms become more sophisticated. These demands have traditionally required local high-performance computing clusters. Although effective, clusters cost millions of dollars, take up considerable space and require numerous support staff, making them infeasible for many fMRI researchers and institutions. Moreover, these systems cannot be flexibly expanded as needs grow rapidly. Cloud computing represents a different kind of solution, providing scalable resources in an affordable and accessible way. Such resources can be harnessed using SaaS, allowing researchers to perform analyses on demand from anywhere in the world without developing and deploying software or managing servers. SaaS has previously been used in neuroscience[100] and has been adopted by several other scientific domains.

Real-time fMRI is a good use case for SaaS because of the need for fast, scalable and resilient computing. Each acquired brain volume would be sent to the cloud for analysis and the result returned before the next volume is finished being collected, for neurofeedback or experiment adjustment. Beyond meeting these tight time deadlines, the goal of real-time cloud analysis is to enable online use of the full breadth of analyses that can be performed offline. As a test case that is particularly intensive computationally, we have been developing a real-time cloud version of FCMA[97]. The computations related to correlation calculation, feature selection and classifier training are especially demanding of memory and compute cycles. An even more demanding factor in a real-time application is that these steps need to be performed multiple times, incrementally as each brain volume is acquired, rather than once at the end from batch data.

In implementing real-time FCMA, we designed a system architecture that can, in principle, be used to perform other intensive fMRI analyses in real-time (Fig. 7). The scanner control room is equipped with a simple workstation that collects and transmits reconstructed brain volumes as they come off the scanner. It uses a hypertext transfer protocol (HTTP) interface to transmit and receive information to and from the cloud. The cloud hosts a representational state transfer (REST) frontend server that communicates with a distributed backend (orchestrated by a master process), which provides a flexible set of processes that compose the stages of the pipeline. Some of the stages involve distributed parallel processing across a set of machines for large computations. Simpler stages, like spatial filtering and classifier scoring, only require a single machine. The system is able to provide multiple services, allowing real-time fMRI experiments from different neuroimaging centers to run simultaneously. The system is also designed to be fault tolerant and robust to machine failures, ensuring its reliability for scientific research. Such cloud services eventually stand to benefit all fMRI analyses, improving standardization and replicability—only a single instance of the backend hardware and software is needed—and making large-scale analyses that once required dedicated systems tractable and accessible to all users.

Real-time fMRI is typically performed on local workstations, even for advanced analyses[47,48], so it is important to consider what benefits SaaS confers beyond enabling particularly large-scale approaches like FCMA. Some practical benefits are mentioned above—increasing the accessibility of real-time fMRI to sites without powerful local workstations, as well as facilitating software configuration, maintenance and upgrades—but there are more general computational advantages as well, related to parallelization. In particular, SaaS allows for the flexible allocation of machines to process the same data in multiple ways. For example, if neurofeedback is to be provided from MVPA, multiple classifiers could be trained and tested on different brain regions or searchlights, and the best-performing classifier could then be used to provide high-fidelity feedback. Likewise, the space of analysis parameters (for example, the penalty in regularized logistic regression) could be swept in real-time to optimize performance. A potential limitation of SaaS is that it involves a central hardware resource that must be supported by fees or major capital investment. Another important consideration is that scanner images can contain identifiable information about human participants (for example, header files, face reconstructions), and thus sending them offsite to locations not explicitly covered by ethics approvals could be problematic. Some cloud computing providers offer HIPAA-compliant servers, which would help alleviate this concern, and additionally, the images could be scrubbed of identifiable information before transmission to the server.

CONCLUSIONS

It is sometimes easy to forget that fMRI has been around for little more than two decades. In this article, we summarized analysis approaches that are helping lead the way into the third decade. Even if our perspective on which directions seem most promising turns out to not be accurate, the underlying theme seems likely to prove correct: fMRI analysis will benefit from close alignment with neighboring fields, such as cognitive science, computer science, engineering, statistics and mathematics. As these fields (and neuroscience) become increasingly represented in the technology industry, new opportunities for funding and partnerships will also arise, along with new applications of the research. In parallel, a growing focus on reproducibility, public databases, code sharing and citizen science promises new, affordable and accessible sources of data and new avenues for discovery. Even if the researchers who collected a dataset find their needle in the haystack, there are surely more needles hidden in it, especially when combined with other datasets and analyzed with powerful techniques being developed.

Acknowledgments

We thank C. Chen, M. Regev, Y. Wang, and H. Zhang for assistance and the members of our labs at Princeton University and Intel Labs for their numerous invaluable contributions to the work described herein. This work was made possible by support from Intel Corporation, the John Templeton Foundation, NIH grants R01 EY021755 and R01 MH069456, and NSF grant MRI BCS1229597. The opinions expressed in this publication do not necessarily reflect the views of these agencies.

References

1. Bandettini PA, Wong EC, Hinks RS, Tikofsky RS, Hyde JS. Time course EPI of human brain function during task activation. Magn Reson Med. 1992; 25:390–397. [PubMed: 1614324]
2. Kwong KK, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA. 1992; 89:5675–5679. [PubMed: 1608978]
3. Ogawa S, et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci USA. 1992; 89:5951–5955. [PubMed: 1631079]
4. Posner MI, Petersen SE, Fox PT, Raichle ME. Localization of cognitive operations in the human brain. Science. 1988; 240:1627–1631. [PubMed: 3289116]
5. Friston KJ, et al. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp. 1994; 2:189–210.
6. Braver TS, et al. A parametric study of prefrontal cortex involvement in human working memory. Neuroimage. 1997; 5:49–62. [PubMed: 9038284]
7. Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp. 1997; 5:329–340. [PubMed: 20408237]
8. Forman SD, et al. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med. 1995; 33:636–647. [PubMed: 7596267]
9. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage. 2009; 44:83–98. [PubMed: 18501637]
10. Ashburner J. SPM: a history. Neuroimage. 2012; 62:791–800. [PubMed: 22023741]
11. Cox RW. AFNI: what a long strange trip it's been. Neuroimage. 2012; 62:743–747. [PubMed: 21889996]
12. Goebel R. BrainVoyager--past, present, future. Neuroimage. 2012; 62:748–756. [PubMed: 22289803]
13. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. Neuroimage. 2012; 62:782–790. [PubMed: 21979382]
14. Lewis-Peacock, JA., Norman, KA. Multi-voxel pattern analysis of fMRI data. In: Gazzaniga, M., Mangun, R., editors. The Cognitive Neurosciences. MIT Press; 2014. p. 911-920.
15. Haxby JV, et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 2001; 293:2425–2430. [PubMed: 11577229]
16. Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. Nat Neurosci. 2005; 8:679–685. [PubMed: 15852014]
17. Freeman J, Brouwer GJ, Heeger DJ, Merriam EP. Orientation decoding depends on maps, not columns. J Neurosci. 2011; 31:4792–4804. [PubMed: 21451017]
18. Kriegeskorte N, Diedrichsen J. Inferring brain-computational mechanisms with models of activity measurements. Phil Trans R Soc B. 2016; 371:20160278. [PubMed: 27574316]
19. Ryali S, Supekar K, Abrams DA, Menon V. Sparse logistic regression for whole-brain classification of fMRI data. Neuroimage. 2010; 51:752–764. [PubMed: 20188193]
20. Etzel JA, Zacks JM, Braver TS. Searchlight analysis: promise, pitfalls, and potential. Neuroimage. 2013; 78:261–269. [PubMed: 23558106]
21. Wimmer GE, Shohamy D. Preference by association: how memory mechanisms in the hippocampus bias decisions. Science. 2012; 338:270–273. [PubMed: 23066083]
22. Zeithamova D, Dominick AL, Preston AR. Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. Neuron. 2012; 75:168–179. [PubMed: 22794270]
23. Kim G, Lewis-Peacock JA, Norman KA, Turk-Browne NB. Pruning of memories by context-based prediction error. Proc Natl Acad Sci USA. 2014; 111:8997–9002. [PubMed: 24889631]

24. Gordon AM, Rissman J, Kiani R, Wagner AD. Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. Cereb Cortex. 2014; 24:3350–3364. [PubMed: 23921785]
25. Hindy NC, Ng FY, Turk-Browne NB. Linking pattern completion in the hippocampus to predictive coding in visual cortex. Nat Neurosci. 2016; 19:665–667. [PubMed: 27065363]
26. Deuker L, et al. Memory consolidation by replay of stimulus-specific neural activity. J Neurosci. 2013; 33:19373–19383. [PubMed: 24305832]
27. St-Laurent M, Abdi H, Buchsbaum BR. Distributed patterns of reactivation predict vividness of recollection. J Cogn Neurosci. 2015; 27:2000–2018. [PubMed: 26102224]
28. Todd MT, Nystrom LE, Cohen JD. Confounds in multivariate pattern analysis: theory and rule representation case study. Neuroimage. 2013; 77:157–165. [PubMed: 23558095]
29. Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci. 2008; 2:4. [PubMed: 19104670]
30. Khaligh-Razavi SM, Kriegeskorte N. Deep supervised, but not unsupervised, models may explain IT cortical representation. PLoS Comput Biol. 2014; 10:e1003915. [PubMed: 25375136]
31. Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C. Hippocampal activity patterns carry information about objects in temporal context. Neuron. 2014; 81:1165–1178. [PubMed: 24607234]
32. Chan SCY, Niv Y, Norman KA. A probability distribution over latent causes, in the orbitofrontal cortex. J Neurosci. 2016; 36:7817–7828. [PubMed: 27466328]
33. Schapiro AC, Kustner LV, Turk-Browne NB. Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr Biol. 2012; 22:1622–1627. [PubMed: 22885059]
34. Hulbert JC, Norman KA. Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. Cereb Cortex. 2015; 25:3994–4008. [PubMed: 25477369]
35. Poppenk J, Norman KA. Briefly cuing memories leads to suppression of their neural representations. J Neurosci. 2014; 34:8010–8020. [PubMed: 24899722]
36. Milivojevic B, Vicente-Grabovetsky A, Doeller CF. Insight reconfigures hippocampal-prefrontal memories. Curr Biol. 2015; 25:821–830. [PubMed: 25728693]
37. Schlichting ML, Mumford JA, Preston AR. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat Commun. 2015; 6:8151. [PubMed: 26303198]
38. Wimber M, Alink A, Charest I, Kriegeskorte N, Anderson MC. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. Nat Neurosci. 2015; 18:582–589. [PubMed: 25774450]
39. Favila SE, Chanales AJH, Kuhl BA. Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. Nat Commun. 2016; 7:11066. [PubMed: 27925613]
40. Davis T, et al. What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. Neuroimage. 2014; 97:271–283. [PubMed: 24768930]
41. Alink, A., Walther, A., Krugliak, A., van den Bosch, JJF., Kriegeskorte, N. Mind the drift - improving sensitivity to fMRI pattern information by accounting for temporal pattern drift. 2015. Preprint at http://biorxiv.org/content/early/2015/12/04/032391
42. LaConte SM. Decoding fMRI brain states in real-time. Neuroimage. 2011; 56:440–454. [PubMed: 20600972]
43. Sulzer J, et al. Real-time fMRI neurofeedback: progress and challenges. Neuroimage. 2013; 76:386–399. [PubMed: 23541800]
44. deCharms RC, et al. Control over brain activation and pain learned by using real-time functional MRI. Proc Natl Acad Sci USA. 2005; 102:18626–18631. [PubMed: 16352728]
45. Schnyer DM, et al. Neurocognitive therapeutics: from concept to application in the treatment of negative attention bias. Biol Mood Anxiety Disord. 2015; 5:1. [PubMed: 25905002]

46. Li X, et al. Volitional reduction of anterior cingulate cortex activity produces decreased cue craving in smoking cessation: a preliminary real-time fMRI study. Addict Biol. 2013; 18:739–748. [PubMed: 22458676]
47. Shibata K, Watanabe T, Sasaki Y, Kawato M. Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. Science. 2011; 334:1413–1415. [PubMed: 22158821]
48. deBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB. Closed-loop training of attention with real-time brain imaging. Nat Neurosci. 2015; 18:470–475. [PubMed: 25664913]
49. Cox RW, Jesmanowicz A, Hyde JS. Real-time functional magnetic resonance imaging. Magn Reson Med. 1995; 33:230–236. [PubMed: 7707914]
50. Goddard N, et al. Online analysis of functional MRI datasets on parallel platforms. J Supercomput. 1997; 11:295–318.
51. Weiskopf N, et al. Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. Neuroimage. 2003; 19:577–586. [PubMed: 12880789]
52. Bray S, Shimojo S, O'Doherty JP. Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. J Neurosci. 2007; 27:7498–7507. [PubMed: 17626211]
53. Ramot M, Grossman S, Friedman D, Malach R. Covert neurofeedback without awareness shapes cortical network spontaneous connectivity. Proc Natl Acad Sci USA. 2016; 113:E2413–E2420. [PubMed: 27071084]
54. Harmelech T, Friedman D, Malach R. Differential magnetic resonance neurofeedback modulations across extrinsic (visual) and intrinsic (default-mode) nodes of the human cortex. J Neurosci. 2015; 35:2588–2595. [PubMed: 25673851]
55. Emmert K, et al. Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: how is brain regulation mediated? Neuroimage. 2016; 124(Pt A):806–812. [PubMed: 26419389]
56. Yoo JJ, et al. When the brain is prepared to learn: enhancing human learning using real-time fMRI. Neuroimage. 2012; 59:846–852. [PubMed: 21821136]
57. Leeds DD, Tarr MJ. A method for real-time visual stimulus selection in the study of cortical object perception. Neuroimage. 2016; 133:529–548. [PubMed: 26973168]
58. Daw ND, Doya K. The computational neurobiology of learning and reward. Curr Opin Neurobiol. 2006; 16:199–204. [PubMed: 16563737]
59. Tom SM, Fox CR, Trepel C, Poldrack RA. The neural basis of loss aversion in decision-making under risk. Science. 2007; 315:515–518. [PubMed: 17255512]
60. Hampton AN, Bossaerts P, O'Doherty JP. Neural correlates of mentalizing-related computations during strategic interactions in humans. Proc Natl Acad Sci USA. 2008; 105:6741–6746. [PubMed: 18427116]
61. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. Neuron. 2011; 69:1204–1215. [PubMed: 21435563]
62. Niv Y, et al. Reinforcement learning in multidimensional environments relies on attention mechanisms. J Neurosci. 2015; 35:8145–8157. [PubMed: 26019331]
63. Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND. Model-based choices involve prospective neural activity. Nat Neurosci. 2015; 18:767–772. [PubMed: 25799041]
64. Boorman ED, Rajendran VG, O'Reilly JX, Behrens TE. Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. Neuron. 2016; 89:1343–1354. [PubMed: 26948895]
65. (Cameron), Chen P-H., et al. A reduced-dimension fMRI shared response model. In: Cortes, C.Lawrence, ND.Lee, DD.Sugiyama, M., Garnett, R., editors. Advances in Neural Information Processing Systems. Vol. 28. Curran Associates, Inc; 2015. p. 460-468.
66. Haxby JV, et al. A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron. 2011; 72:404–416. [PubMed: 22017997]
67. Wang, X., Hutchinson, R., Mitchell, TM. Thrun, S.Saul, LK., Schölkopf, PB., editors. Training fMRI classifiers to discriminate cognitive states across multiple subjects. Advances in Neural

Information Processing Systems 16. 2003. https://papers.nips.cc/paper/2449-training-fmri-classifiers-to-detect-cognitive-states-across-multiple-human-subjects

68. Richter FR, Chanales AJH, Kuhl BA. Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. Neuroimage. 2016; 124:323–335. [PubMed: 26327243]
69. Poldrack RA, Halchenko YO, Hanson SJ. Decoding the large-scale structure of brain function by classifying mental States across individuals. Psychol Sci. 2009; 20:1364–1372. [PubMed: 19883493]
70. Finn ES, et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci. 2015; 18:1664–1671. [PubMed: 26457551]
71. Guntupalli JS, et al. A model of representational spaces in human cortex. Cereb Cortex. 2016; 26:2919–2934. [PubMed: 26980615]
72. Chen J, et al. Shared memories reveal shared structure in neural activity across individuals. Nat Neurosci. 2017; 20:115–125. [PubMed: 27918531]
73. Dubois J, Adolphs R. Building a science of individual differences from fMRI. Trends Cogn Sci. 2016; 20:425–443. [PubMed: 27138646]
74. Rosenberg MD, et al. A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci. 2016; 19:165–171. [PubMed: 26595653]
75. Manning JR, Ranganath R, Norman KA, Blei DM. Topographic factor analysis: a Bayesian model for inferring brain networks from neural data. PLoS One. 2014; 9:e94914. [PubMed: 24804795]
76. Turk-Browne NB. Functional interactions as big data in the human brain. Science. 2013; 342:580–584. [PubMed: 24179218]
77. Al-Aidroos N, Said CP, Turk-Browne NB. Top-down attention switches coupling between low-level and high-level areas of human visual cortex. Proc Natl Acad Sci USA. 2012; 109:14675–14680. [PubMed: 22908274]
78. Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex. 2012; 22:158–165. [PubMed: 21616982]
79. Cole MW, et al. Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Neurosci. 2013; 16:1348–1355. [PubMed: 23892552]
80. Fornito A, Zalesky A, Bullmore ET. Network scaling effects in graph analytic studies of human resting-state fMRI data. Front Syst Neurosci. 2010; 4:22. [PubMed: 20592949]
81. Wang Y, Cohen JD, Li K, Turk-Browne NB. Full correlation matrix analysis (FCMA): an unbiased method for task-related functional connectivity. J Neurosci Methods. 2015; 251:108–119. [PubMed: 26004849]
82. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009; 10:186–198. [PubMed: 19190637]
83. Giusti, C., Ghrist, R., Bassett, DS. Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data. 2016. Preprint at https://arxiv.org/abs/1601.01704
84. Mitchell TM, et al. Predicting human brain activity associated with the meanings of nouns. Science. 2008; 320:1191–1195. [PubMed: 18511683]
85. Anderson JR. A spreading activation theory of memory. J Verbal Learn Verbal Behav. 1983; 22:261–295.
86. Cohen JD, Dunbar K, McClelland JL. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychol Rev. 1990; 97:332–361. [PubMed: 2200075]
87. Yamins DLK, DiCarlo JJ. Using goal-driven deep learning models to understand sensory cortex. Nat Neurosci. 2016; 19:356–365. [PubMed: 26906502]
88. Kriegeskorte N. Deep neural networks: a new framework for modeling biological vision and brain information processing. Annu Rev Vis Sci. 2015; 1:417–446. [PubMed: 28532370]
89. Cichy, RM., Khosla, A., Pantazis, D., Torralba, A., Oliva, A. Deep neural networks predict hierarchical spatio-temporal cortical dynamics of human visual object recognition. 2016. Preprint at https://arxiv.org/abs/1601.02970

90. Güçlü U, van Gerven MAJ. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. J Neurosci. 2015; 35:10005–10014. [PubMed: 26157000]

91. Anderson JR. Tracking problem solving by multivariate pattern analysis and Hidden Markov Model algorithms. Neuropsychologia. 2012; 50:487–498. [PubMed: 21820455]

92. Wang, Y., et al. Full correlation matrix analysis of fMRI data on Intel Xeon Phi coprocessors. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (Supercomputing); ACM; 2015.

93. Anderson, MJ., et al. Enabling factor analysis on thousand-subject neuroimaging datasets. 2016. Preprint at https://arxiv.org/abs/1608.04647

94. Dean, J., et al. Large scale distributed deep networks. In: Pereira, F.Burges, CJC.Bottou, L., Weinberger, KQ., editors. Advances in Neural Information Processing Systems 25. Curran Associates, Inc; 2012. p. 1223-1231.

95. Gorgolewski KJ, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data. 2016; 3:160044. [PubMed: 27326542]

96. Simony E, et al. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat Commun. 2016; 7:12141. [PubMed: 27424918]

97. Wang, Y., et al. Real-time full correlation matrix analysis of fMRI data; 2016 IEEE International Conference on Big Data (Big Data 2016); Curran Associates, Inc; in the press

98. Vodrahalli, K., et al. Mapping between natural movie fMRI responses and word-sequence representations. 2016. Preprint at https://arxiv.org/abs/1610.03914

99. Zhang, H., et al. A searchlight factor model approach for locating shared information in multi-subject fMRI analysis. 2016. Preprint at https://arxiv.org/abs/1609.09432

100. Freeman J, et al. Mapping brain activity at scale with cluster computing. Nat Methods. 2014; 11:941–950. [PubMed: 25068736]

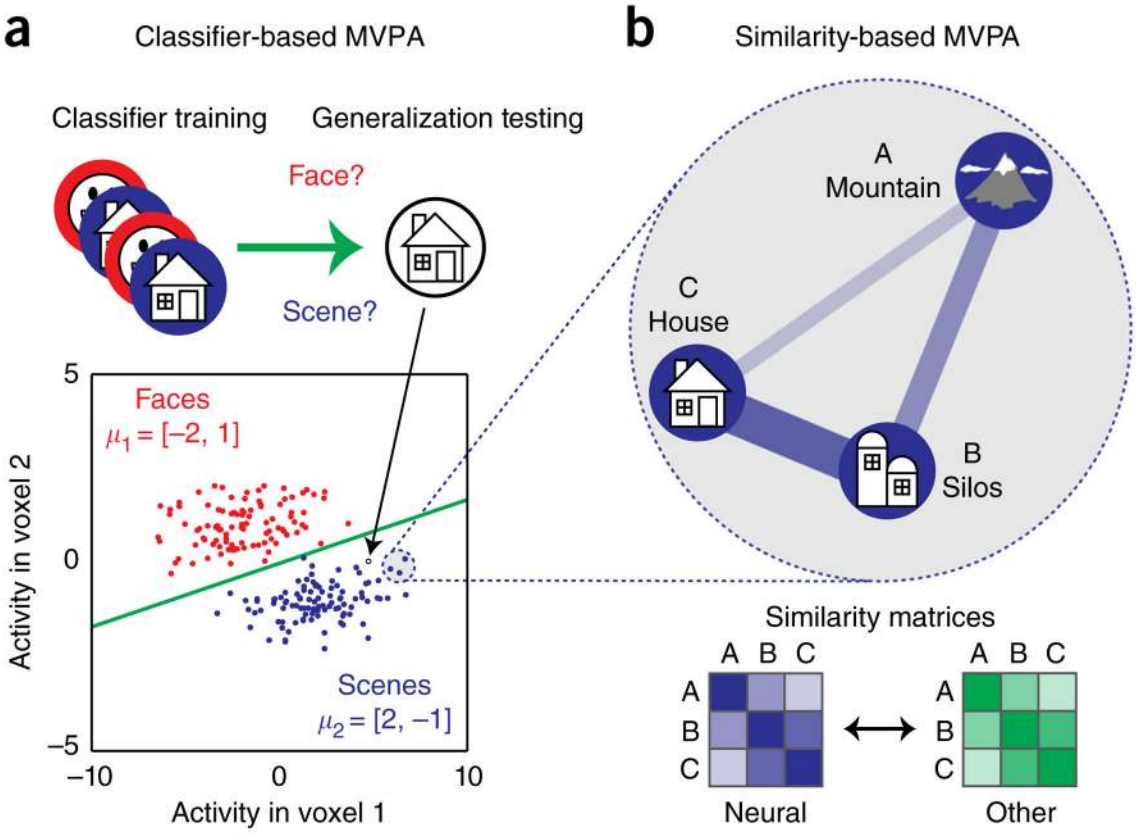

Figure 1.

Types of MVPA. (**a**) Classifier-based MVPA involves learning a boundary that discriminates between fMRI patterns associated with different cognitive states (for example, attending to faces vs. scenes). (**b**) Similarity-based MVPA involves computing the matrix of pairwise distances between fMRI patterns and (optionally) comparing this matrix to other similarity matrices (for example, predictions from a cognitive theory about conceptual similarity). Adapted with permission from ref. 14, J.A. Lewis-Peacock and K.A. Norman, in *The Cognitive Neurosciences, fifth edition,* edited by Michael S. Gazzaniga and George R. Mangun, published by The MIT Press.

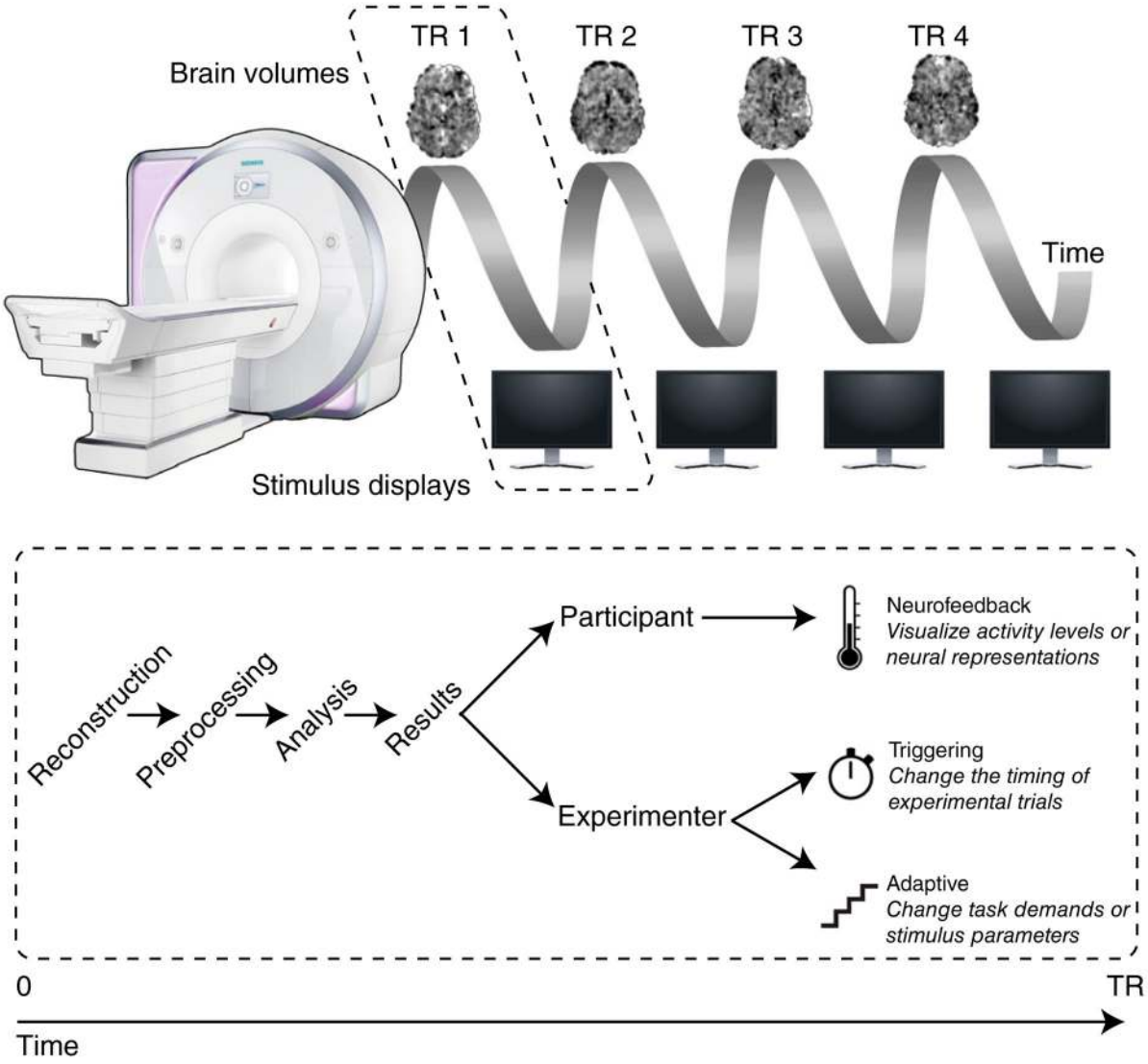

Figure 2.

Uses of real-time fMRI. After a brain volume is collected, preprocessed and analyzed, the results are used to update the experimental code before the next acquisition is completed. The neural results can be incorporated in three general ways. First, they can be displayed directly to participants using a scale, gauge or an aspect of the stimulus itself, with the idea that this neurofeedback will allow participants to refine their strategies and learn to control which brain regions or representations are active. Second, the results can be used by the experimenter to trigger the onset of a trial, to test hypotheses about the contribution of a region's activity (or inactivity) to a cognitive process or behavior of interest. Third, the results can be used by the experimenter to adjust experimental parameters, either to hone in on the selectivity of brain regions or to repeat trials or conditions for which there is uncertainty about the neural response. Regardless of the approach, the end result is that brain activity at one timestep can influence the participant's experience at the next timestep, which in turn influences their brain activity, then their experience and so on. If the brain and

experiment are tightly entrained, the design can be referred to as ‘closed-loop’. TR refers to repetition time, the start of the next volume acquisition.

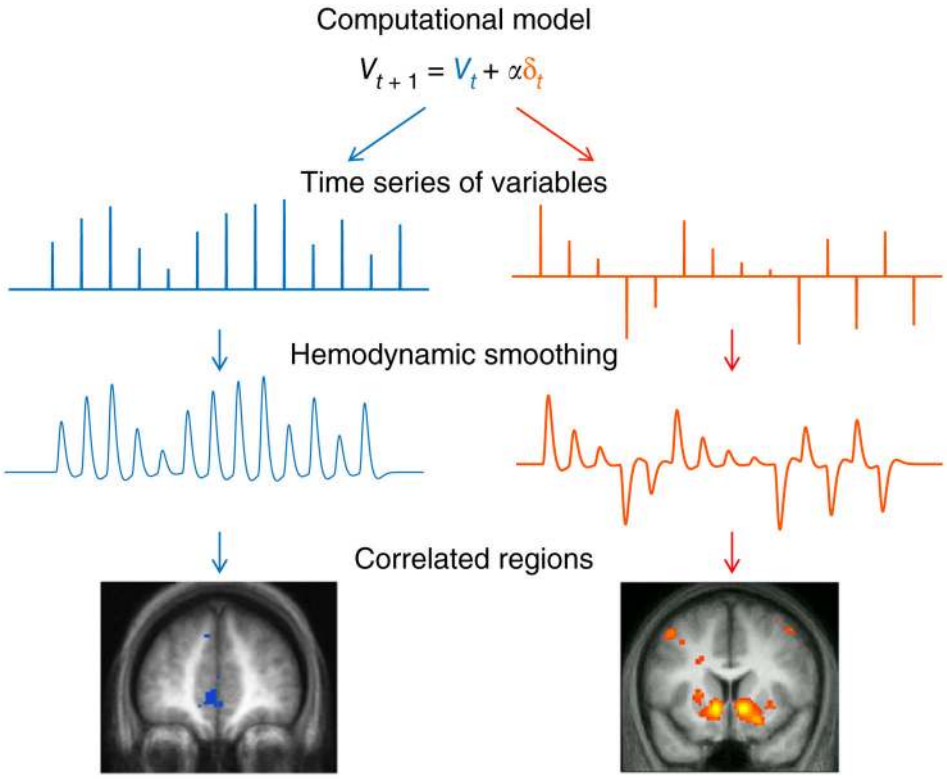

Figure 3.

Schematic of model-based fMRI. A computational model (top; error-driven reward-learning model) can be used to generate candidate time series of internal variables (values, V, in blue and prediction errors, δ, in red). Following smoothing to account for the hemodynamic lag (middle), these variables can be used as regressors to seek correlated BOLD activity in the brain (bottom), producing regions that can be considered candidates for performing or tracking the corresponding computations in the model. α, learning rate parameter.

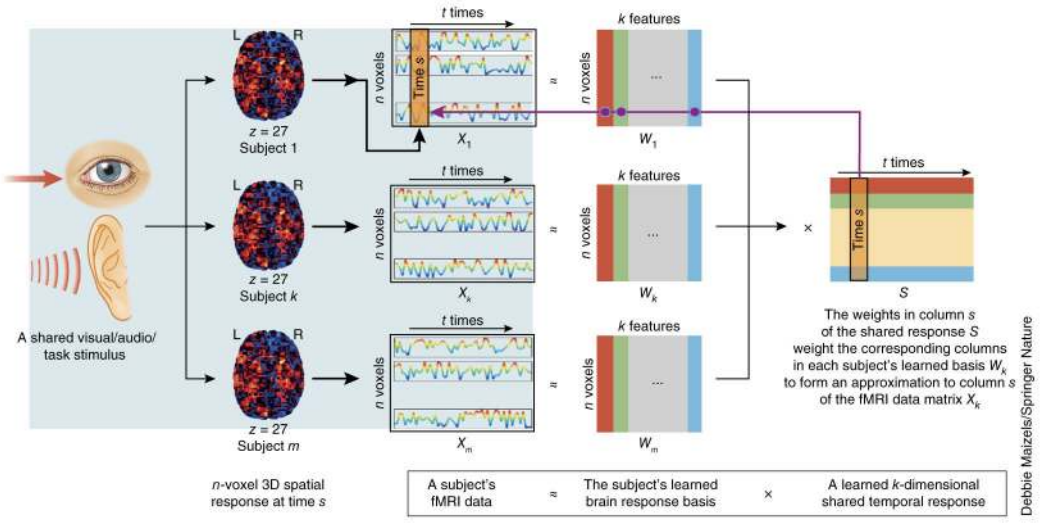

Figure 4.

Process for SRM. fMRI data are collected from each of m participants experiencing the same stimulus and then organized into a matrix X (voxels by time). Each matrix X is then factored using a probabilistic latent-factor model into the product of a subject-specific matrix W of k brain maps (an orthogonal basis) and a shared temporal response matrix S of size k by time. That is, for each participant: $X = WS + R$, where X, W, and the residuals, R (not shown), are subject-specific, and S is shared across participants.

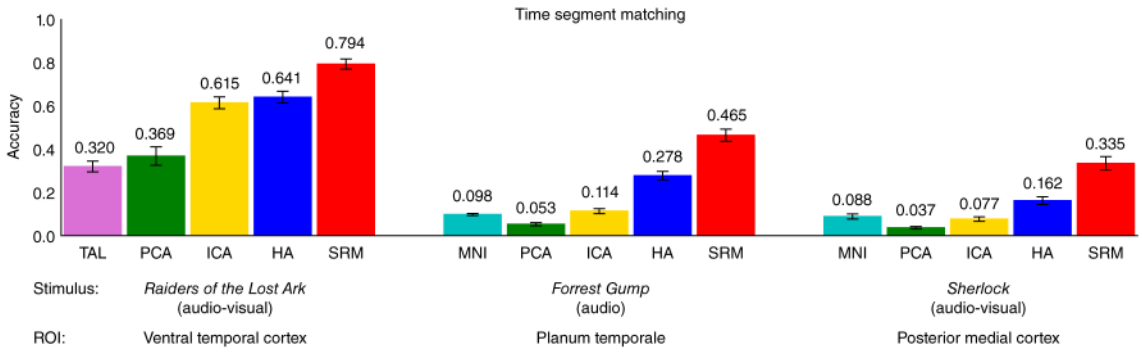

Figure 5.

Comparison of SRM to other multisubject approaches. Three fMRI datasets were collected as participants viewed and/or listened to the same stimulus. The data were then anatomically aligned to either Talairach (TAL) or Montreal Neurological Institute (MNI) space, giving a common coordinate system for all participants. In these examples, analyses were restricted to the time series of BOLD activity across voxels within a specific ROI. The strength of shared cognitive states during the stimulus was evaluated by attempting to identify a short movie segment in a held-out participant's test data, based on the test data of the other participants. This was done before (TAL, MNI) or after applying an across-participant factor model: principal component analysis (PCA), independent component analysis (ICA), hyperalignment (HA) or SRM. For the methods with dimensionality reduction (PCA, ICA, SRM), $k = 50$ features were used. Assuming the movie segments are independent, chance accuracy is 0.001. Error bars reflect s.e.m. Adapted with permission from ref. 65, Curran Associates, Inc.

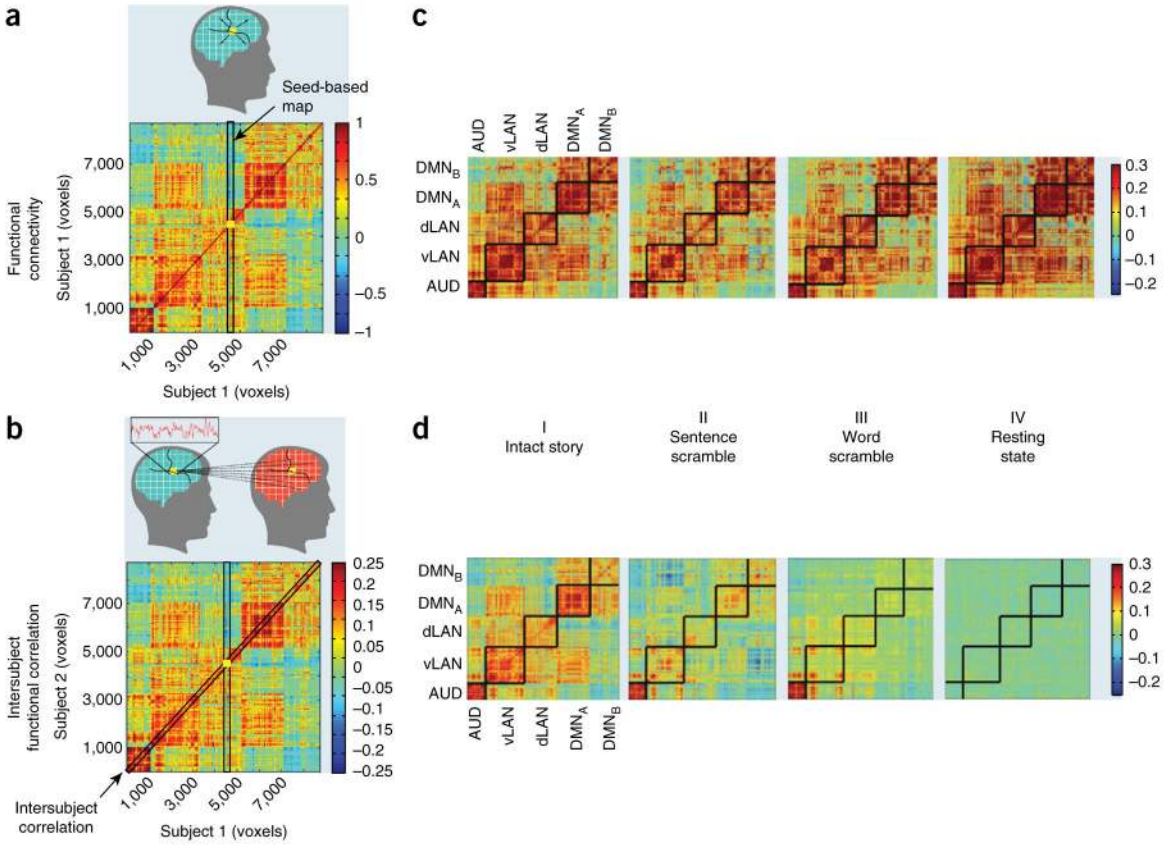

Figure 6.

Functional interactions within and between participants. (**a**) Schematic of functional connectivity (FC) analysis, performed within a participant's brain between a single seed ROI (in yellow) and the rest of the brain or between all possible pairs of voxels (as in FCMA). (**b**) Schematic of intersubject functional correlation (ISFC) analysis, performed between participants' brains from a single seed ROI to the rest of the brain or between all possible pairs of voxels. (**c**) Voxel-based FC covariance matrices reveal similar network organization across conditions. The four experimental conditions are: (i) intact story, (ii) sentence scramble, (iii) word scramble and (iv) resting state. (**d**) Voxel-based ISFC covariance matrices reveal stimulus-dependent interactions within networks. The covariance matrices are organized into five networks by *k*-means clustering: default-mode network, split into two subnetworks (DMN_A and DMN_B); dorsal language network (dLAN); ventral language network (vLAN); and auditory cortex (AUD). Adapted with permission from ref. 96, Nature Partner Journals.

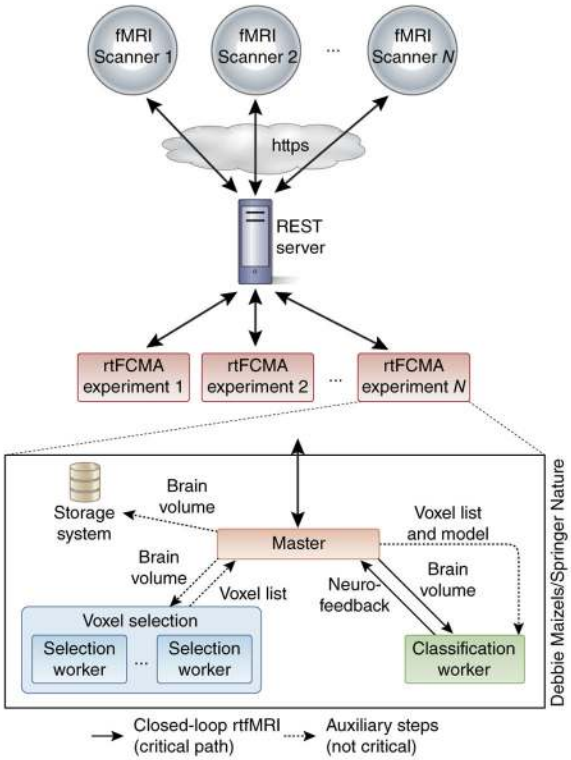

Figure 7.

Real-time cloud system architecture. Scanners send brain volumes to the system via the Internet and the system returns neurofeedback. Using a series of servers, each scanner gets its own instance of the analysis pipeline (here, FCMA), providing neurofeedback in real-time and keeping the classifier updated by performing feature selection and classifier training iteratively. Adapted with permission from ref. 97, IEEE.