Exploratory Data Analysis Methods for Functional Magnetic Resonance Imaging (fMRI): A Comprehensive Review of Software Programs Used in Research

Hussain A. Jaber[1*], Basma A. Al-Ghali[2], Muna M. Kareem[3], Ilyas Çankaya[4], Oktay Algin[5]

Authors affiliations:

1*) Department of Biomedical Engineering, College of Engineering, Al-Nahrain University, Baghdad, Iraq.
hussainjaber2000@nahrainuniv.edu.iq

2) Department of Biomedical Engineering, College of Engineering, Al-Nahrain University, Baghdad, Iraq.
basma.a.faihan@nahrainuniv.edu.iq

3) Department of Biomedical Engineering, College of Engineering, Al-Nahrain University, Baghdad, Iraq.
Muna.kareem@nahrainuniv.edu.iq

4) Electrical and Electronics Engineering Department, Graduate School of Natural Science, Ankara Yıldırım Beyazıt University, Ankara, Turkey
ilyas.cankaya@aybu.edu.tr

5) Interventional MR Clinical R&D Institute, Ankara University, Ankara, Turkey
Department of Radiology, Medical Faculty, Ankara University, Ankara, Turkey
National MR Research Center (UMRAM), Bilkent University, Ankara, Turkey.
oktay.algin@umram.bilkent.edu.tr

Paper History:

Received: 26th Sep. 2024

Revised: 26th Sep. 2024

Accepted: 12th Oct. 2024

Abstract

This extensive and thorough review aims to systematically outline, clarify, and examine the numerous exploratory data analysis techniques that are employed in the intriguing and rapidly advancing domain of functional MRI research. We will particularly focus on the wide array of software applications that are instrumental in facilitating and improving these complex and often nuanced analyses. Throughout this discourse, we will meticulously assess the various strengths and limitations associated with each analytical tool, offering invaluable insights relevant to their application and overall efficacy across diverse research contexts and environments. Our aim is to create a comprehensive understanding of how these tools can be best utilized to enhance research outcomes. Through this analysis, we aspire to equip researchers with critical knowledge and essential information that could profoundly influence their methodological selections in upcoming studies. By carefully considering these factors, we hope to contribute positively to the ongoing progression of this important field of inquiry, fostering innovation and enhancing the impact of future research findings in functional MRI studies.

Keywords: Functional MRI, Machine Learning, Data Analysis, Brain Imaging, Brain Mapping, Neuroimaging Analysis Software, Statistical Parametric Mapping (SPM).

استكشاف طرق تحليل بيانات التصوير بالرنين المغناطيسي fMRI : مراجعة شاملة لادوات البرامج المستخدمة في البحث

حسين عبد جابر، بسمة عبد الصاحب فيحان، منى مصطفى كريم، Ilyas Çankaya, Oktay Algin

الخلاصة:

هدف هذا العمل الشامل والموسع إلى تحديد وتوضيح وفحص العديد من تقنيات تحليل بيانات التصوير بالرنين المغناطيسي fMRI والمستخدمة في تحليل بيانات علم الأعصاب المثير للاهتمام والسريع التقدم. وسوف نركز بشكل خاص على مجموعة واسعة من تطبيقات البرامج التي تساعد في تسهيل وتحسين هذه التحليلات المعقدة والدقيقة في كثير من الأحيان. وخلال هذا البحث، سوف نقوم بتقييم نقاط القوة والقيود المختلفة المرتبطة بكل أداة تحليلية بدقة، وتقديم رؤى لا تقدر بثمن ذات صلة بتطبيقها وفعاليتها الشاملة عبر سياقات وبيئات بحثية متنوعة. هدفنا هو خلق فهم شامل لكيفية الاستفادة من هذه الأدوات على أفضل وجه لتعزيز نتائج البحث. ومن خلال هذا التحليل، نطمح إلى تزويد الباحثين بالمعرفة الحاسمة والمعلومات الأساسية التي يمكن أن تؤثر بشكل عميق على اختياراتهم المنهجية في الدراسات القادمة. ومن خلال النظر بعناية في هذه العوامل، نأمل أن نساهم بشكل إيجابي في التقدم المستمر لهذا المجال المهم من الاستقصاء، وتعزيز الابتكار وتعزيز تأثير نتائج الأبحاث المستقبلية في دراسات التصوير بالرنين المغناطيسي الوظيفي.

1. Introduction

In recent years, functional magnetic resonance imaging (fMRI) has developed into an essential approach for examining the neural mechanisms that underpin human cognition as well as various medical conditions in a non-invasive manner. With the rising utilization of this technique within the domains of social and cognitive neuroscience, the demand for dependable and transparent analytical methods has become increasingly significant. [1] This paper offers a comprehensive review of the frequently employed exploratory data analysis techniques in fMRI research. The aim is to provide readers with an essential understanding of the current landscape of the field, incorporating a detailed assessment of the methodologies employed by researchers and the software tools that have gained traction among professionals in this discipline.

The specific objectives of the review were a) to conduct a scoping review and author survey of the literature to identify the leading software programs used for exploratory data analysis in functional MRI research, the target journals and impact factors that publish studies utilizing these techniques, as well as the details of the authors conducting them, b) based on the review, discuss the exploration tools and methodologies, with the relevant software programs available for neuroimaging analysis, and c) conduct a secondary literature search to identify software programs that are not necessarily part of the classification under review. The use of neuroimaging in both cognitive and clinical research is becoming increasingly more accessible. With more data comes a greater responsibility for those working at the bench; primarily to understand not only the more complex region-specific aspects of function, but also how these processes are integrated across various networks of brain regions. [2, 3]

Equally important, when studying the multi-faceted complexities of brain function, is an effective technique for exploratory analysis that helps consolidate the many details of experiment design—be they region-specific or widely distributed networks. This is why many researchers, dealers in tools like the General Linear Model, might appreciate a paper that outlines the main techniques and tools with which to cluster regions into related sets. Given the growing interest in the field and the need for a technique that helps to consolidate the kind of details needed to make plethoric multi-faceted inferences about brain function, it is the aim of this work to present a detailed look at the many different approaches and software programs that are currently in use. Readers new to the field might begin by reading through Sections "State of Play", "Author Details" and "Methods", before following up with the specifics of the software programs and methodologies in "Exploration." Even though they did not fit into this kind of review paper, they are valuable tools that those entering the field might also find useful. [4][5]

2. Functional MRI: Overview and Data Collection

Functional Magnetic Resonance Imaging (MRI) is a non-invasive neuroimaging method that measures brain activity indirectly via the dynamic changes in blood flow or blood oxygenation. The basal principle that governs fMRI lies in the coupling between neuronal activity and regional brain hemodynamics, which could lead to two types of contrast in the MR images of the brain. [1] The first type is T2*-weighted images, where the Blood Oxygen Level Dependent (BOLD) contrast is responsible for the susceptibility effects of deoxygenated blood, and the second type is Arterial Spin Labeling (ASL), which reflects flow-related contrast and does not utilize contrast agents. Image data utilized in fMRI is collected from MRI scanners, which could be of the clinical type or the research-dedicated ones. The settings of MRI scanners used to collect fMRI could range in many technical details. Typically, field strengths of 3T and above are preferred for fMRI studies to provide a higher signal-to-noise ratio and to integrate a time series of 3D volumes. Below 3T, statistical power decreases but provides the ability to offer smooth correction methods. [6-8]

Acquiring spatial in fMRI is higher in quality at around 1 mm isotropic voxels, but could go below 1 mm voxel size. BOLD non-linearity increases in such fine measurements, which could result in tissue folding miscommunications. Racevg-White contrast also increases greater amounts when 3 mm spatial resolution is used. At higher temporal resolution, the measurement refresh time should be under 2 seconds to capture the BOLD hemodynamic rate. [9][10] The pre-imaging acquisition of subjects, interventional and experimental designs, are main contributors to high-quality data collection. fMRI is a sensitive method reflecting the stimulus parameters defined in the experiment; hence, a low signal-to-noise ratio with high bias in the data will make the data unreliable and sensitive to data analysis methods. Data misfit in the experiment would enhance the contribution of noise and artifact mechanisms in the data. Artifacts could be in the form of haze or drastic changes in time series. Also, the location of artifacts in the brain could display patterns of influence by the initial placement and removal of the artifact. Therefore, the choice of data collection criteria and quality control at data collection prevents many difficulties in the data analysis, and it is recommended in all works of fMRI resting-state. [11-13]

3. Exploratory Data Analysis in Functional MRI

Functional magnetic resonance imaging (fMRI) yields data with a complex structure often driven by time-dependent dynamics from the blood oxygen level-dependent (BOLD) signal. Today's analyses perform many levels of denoising to untangle these structures. In part, these complexities make fMRI data hard to analyze using statistical modeling alone. Exploratory data analysis (EDA) is performed on fMRI data, and it precedes formal statistical analyses. The goal is to uncover any patterns in the data, and these analyses can be conducted using graphical as well as quantitative methods. The results from this step can guide formal modeling and serve as a data quality check in light of the research question. Over the course of an analysis, the boundary between EDA and formal modeling can shift as more characteristics of the data are discovered. [14][15][16]

Quantitative measures are often developed hand-in-hand with formal statistical techniques. In many cases of complex data structures, these techniques are the only feasible attempt to arrive at the formal statistical properties. In this review, we articulate common EDA methods and guide the reader toward useful software tools for conducting analysis in this sphere. We encourage careful thinking about EDA and visual inspection before any formal statistics are applied. All data have limitations; however, spending time on purely exploratory analyses should be guided by the research question(s) and the data at hand. The exploratory and confirmatory processes progressively employ graphical and quantitative characterization, respectively. A careful storyline is offered in the lifecycle of a study performing EDA, ensuring that scientists and readers are aware of the choices made for data and statistical models. [17][18]

3.1 Preprocessing Steps

Preprocessing plays a crucial role in the targets of any neuroimaging study. In functional MRI, which generates a lot of noise from a vast amount of temporal data, many preprocessing methods have been developed as effective data analyses. One of the most common tools used in fMRI preprocessing is motion correction. The amount of head motion can be marked as a potential outcome of standard exploratory analysis and included as a possible confound in more detailed analysis. Large motion events, or higher mean levels of head motion across the sample of individuals, can have a confounding impact on a number of major measurements in exploratory fMRI data analysis, including seed-based correlation analysis and ICA analysis. Slice timing correction is also required for correcting the temporal differences in the acquisition time. [19, 20]

When preprocessing functional scans, another routine choice to consider is spatial normalization. Such options will require consideration of the gradient descent rate, the regularization coefficient, the number of modeling intensity histograms, the number of nonlinear iteration steps, and the option to estimate and regularize the Jacobian determinants. The ability of different preprocessing steps to change the topography of findings has been shown at the level of task- and resting-state fMRI-based connectivity. Experts have noted that changes in preprocessing methods can change the results of clinical prediction studies. This shows that large MRI projects need to make careful deliberation about what preprocessing methods to choose and should be prepared to ascertain how preprocessing impacts their results if exploratory data analysis is involved in their development plan. In summary, careful preprocessing can help ensure a high level of data quality and that the emerging images of brain function remain influential. The exploratory data analysis will help provide useful guidance on which data to validate and incorporate into the next process in the data pipeline. [21, 22]

3.2 Statistical Analysis Techniques

In the paradigms of both psychological and cognitive neuroscience, the primary objectives are the inference of cognitive functioning and the estimation of neurological relationships involved in processing. Therefore, the majority of statistical analysis techniques proposed for fMRI studies are related to revealing these hidden relationships and detecting patterns in brain activities, based on the correspondence between brain functions and cognitive processes. The following subsections present the various statistical analyses, including multiple comparison corrections, that can be performed using sMRI and fMRI findings, and describe a few examples to demonstrate the statistical results. The exploration is performed based on the following relationships established for the associated datasets: [23, 24]

- Cross-sectional studies: A single statistical approach for examining the differences across the results of the control and patient groups.
- Longitudinal studies: Various statistical strategies for comparing the changes across the results of the pre- and post-interventions, including the control and experimental groups.
- Treatment effect: Various regression models that identify the relationships of digitized values with the associated response variables for semi-quantification studies. [25, 26]

Parametric and non-parametric statistics are the most common statistical tools used in sMRI and fMRI analysis. Therefore, it is necessary to choose statistics that best arise from the study designs and objectives. A high understanding of statistics and image processing could automatically ensure that the statistical assumptions, errors, and biases that could potentially be addressed in the underlying statistical methods used are well understood. Indeed, combined with the multiple comparison corrections to obtain the corrected statistics, this is the key evidence presented to perform valid neuroimaging findings in detecting functional activations. The exploration of the findings can be understood using exploratory data analysis and statistical testing. Therefore, these methods applied to fMRI studies would ensure obtaining valid inferences in detecting neurological events. [27, 28]

4 Software Programs for Functional MRI Analysis

A wide range of freely available and commercial software for the analysis of functional magnetic resonance imaging data exists. All software have particular strengths and weaknesses. The following come from mind maps developed during a literature search of functional and resting state functional software that can analyze fMRI. The primary criterion for inclusion is observable use of TVA, BMS, Non-Parametric Method, and VBM. Information particularly about aspects of the software was also included. Subsets of non-human primate software are also included where it was indicated as such. Though spatial registration and normalization tools are part of many programs, specific image-based programs can be found in the image analysis portion of the Aims and Goals of Projects section. [29, 30]

Software is used for the analysis and/or visualization of functional MRI data. Data may be from clinical patient studies, basic imaging studies with the goal to determine functional localization or organization, or population-based functional imaging studies. Note that though specific functions of software are listed, software may offer functions not stated as well. All functional MRI studies analyzed clinical MRI research that used available software can view and be part of these maps. Versions are not current but may provide general information about software when available. For viewing purposes, the embedded image is limited to 2000 pixels. Original images and drawings are printed to the size they were created, up to 18 x 10 inches at 1200 dpi. Dimensions allowable for inclusion and format for the original figure will need to be reviewed. [31]

4.1 Popular Software Programs

For neuroimaging research, SPM is an extremely popular MATLAB-based analysis brain imaging software package. It is regularly updated, so the newest version should be used for better efficiency and functionality improvements. The data format is very robust, thus making it very popular with collaborators. Furthermore, SPM includes a useful interactive tool, termed SPM; it is extremely useful for quickly viewing the brain results. Additionally, exceedingly useful information can be easily obtained and output for the results map when using XJView. Other useful software programs for fMRI data analysis are FSL and AFNI. FSL provides not only a complete analysis but also pathological models, so it can be used for the development of algorithms for other purposes. This flexibility of FSL provides a measure to accelerate further developments. [32, 33]

Resting-state fMRI can be analyzed using these brain imaging software packages as well as more specialized software programs. The SPM for fMRI, CONN, allows the user to apply graphs in fMRI data by utilizing positive correlations, negative correlations, or the summed point-to-point imaging value of each fMRI dataset. However, the data processing options of CONN are somewhat limited compared to another software package, recently Connectome Workbench. This approach would benefit from the user retaining the scripts used to perform data processing with various options. In addition to these specialized software packages, another highly specialized software program, known as REST, is widely used, as it poses several attractive features. [34, 35]

4.2 SPM (Statistical Parametric Mapping)

One of the most frequently used software programs, SPM12, is part of a well-known software platform for neuroimaging organization. At the time of this writing, the most recent version is SPM12 revision 6906. It is constructed on a foundation of MATLAB, the powerful scientific and numerical computing environment with specialized toolboxes for signal processing, statistics, time series analysis, regression analysis, and artificial intelligence used primarily for fMRI data. The pre-processing stages in SPM12 include realignment, normalization, and spatial smoothing of the functional data. One inferior feature, however, is the need for the MATLAB software to operate SPM12. While obtaining MATLAB is free, purchasing a suitable license and storing it on a powerful PC or cloud facility incurs expenses. Many large labs have access to this software and store it on high-speed PCs; however, using SPM12 could be feasible for most users. [36-40]

SPM has a modular structure that includes the core implementation of its theory. This enables users to combine configurations of sequential stages, evaluated in a Bayesian or frequentist methodology, to achieve desired operating results. Data distribution describes effects at each voxel after hypothesis-driven modeling, smoothing, and normalizing procedures. [41, 42] In addition, SPM12 includes classical optimization algorithms like the expectation-maximization and variational Bayes methods. While several advanced methods are offered, it is associated with a major disadvantage; because of its dense core of MATLAB, interpreting and visualizing results in SPM12 may become laborious when an experimental study is complex. To perform second-level analyses that are commonly implemented, additional use of programs like a computation toolbox for fMRI resting-state research is needed. An operating system running more effective and user-friendly SPM-based preprocessing toolboxes than SPM12 had not emerged at the time of our research. Nevertheless, because SPM12 relies on the same fMRI principle of activation to build a general linear model across multiple subjects, compared to other software, the average user may obtain higher significance levels with minimal training. [43-45]

4.3 FSL (FMRIB Software Library)

FSL is a comprehensive library and suite of tools for image analysis of mainly fMRI data, but also to some extent structural brain images. Coordinate-based meta-analysis and algorithm-level analysis for activation detection are two of its main features. The 'FEAT (fMRI Expert Analysis Tool)' is the most commonly used analysis pipeline for model-guided fMRI analysis. This pipeline has a currently developed version that contains functions to pre-process, model, and post-process data to be analyzed, allowing very flexible models and multiple transform inputs based on the images or design structure. To accommodate the common modeling of fMRI data, the pipeline first applies artifact detection for the removal of nuisance variables and non-physiological components within the model for activation effects. The artifact detection functionalities are essential in pipeline analysis, and it has the option to compute motion parameters for confounding removal. [46-48]

The independent component analysis tool and function first use a two-stage PCA-based mean group ICA to get functional connectivity based on functional data. Then, the dual-regression method is used per subject, and FSL's methods are applied for a final group-level analysis. The 'post-stats' post-processing script is also used to enhance results at stage three. A tool for sensitivity mapping of wrapped phase processing is used under the Physiological Noise Modelling Group. FSL's approach to DICCCOL enables analysis of individual and group differences with higher accuracy. It is used in conjunction with DBM for brain map consolidation and investigation of pathway disruptions or damages. R/rana is employed for non-linear registration between a pair space in the Improved Linear Model of Anatomic Normalization. BrEd in FSL is a visualization tool developed for inspecting NIfTI images with different Lightwave buttons for image manipulation like zoom, window, and sending to the overlay manager. [49-51]

FSL and Python libraries also interface with Python. This allows execution of FSL on a map with a script without opening the FSL graphical user interface for projects. Tools to achieve clustering, calls to processing scripts, and general statistical operations are available as command-line functions, allowing FSL to have command line functional capabilities for use in shell scripts. Many hundreds of filter-based regression sophistication algorithms can be embedded within FSLpipe so that FSL's various packages contain extensive toolchains. The use of FSL in the analysis of fMRI data is done in extensive current research due to user-friendliness and data documented outputs. [52-54]

4.4 AFNI (Analysis of Functional NeuroImages)

AFNI is a powerful software package used for the analysis and display of MR neuroimaging data. AFNI is one of the two major software systems that is publicly available for task-based fMRI data analysis. AFNI implements a full graphical programming environment for processing, analyzing, and exploring complex fMRI data. In addition to being able to compute basic models and perform group analyses on fMRI data, many other powerful tools are present in the software, like machine learning and parcellation programs, hypnotic plugins for the combination of task and resting-state studies, as well as tracking and viewing fibers. [11,55,56]

AFNI provides a multi-modal, multi-program collection and processing with open-source software. AFNI has a project data structure for organizing interactive software tools for processing and viewing complex volumetric, time series, and morphometric neuroimaging data. This capability results from an integrated suite of segmentation, registration, surface extraction, manipulation, editing, and combination techniques covering cortex, thalamus, neocortex, etc., and handling most available human and other imaging species. In AFNI, tools for fMRI time series quality control are provided. This tool has a method for time point exclusion from consideration due to environmental motion confounds. These tools assist the operator in ensuring data quality as well as in discriminating effects of interest from artifacts due to motion or other noise sources. [57-59]

4.5 BrainVoyager

BrainVoyager is currently at version 20.2. Comprehensive and integrated tools for preprocessing and statistical analysis are available, including slice-time correction, motion correction, linear trend removal, temporal and spatial preprocessing, and statistical tools. Statistical tools can be extended by adding options of GLM, LSA, and General Linear Modeling. BrainVoyager offers additional powerful tools such as Dynamic, 3D, and 4D Recon, which can visualize functional voxels at multiple time points, generate 3D volumes, and combine multiple 3D or 4D time points into one 3D or 4D surface volume. Digital brain architecture, surface morphology, and other information within the sphere each report on the specific location within the brain. It can also curate the mathematical data for the grand average between the temporal slices. In addition, it provides support for popular surface-mapping techniques such as 2D projection, volume scanning, and top and depth leveling. Furthermore, BrainVoyager also provides a quantitative method to measure results and generate reports, and integrates various data and functional visualization tools for interactive analysis of data. [60, 61]

However, although BrainVoyager has comprehensive tools and supports a wide range of studies, users must purchase it at a higher price and then receive short-run updates and long-term support after the first purchase. Furthermore, flexible downstream post-processing is limited, while training and learning BrainVoyager in data processing and research are still in the initial stage. The following is the use of BrainVoyager, which includes three steps: specifically designed data sequences, spell correction, and statistical calculation, down to the top of the preprocessing step, where the functional MRI data sequence must be organized in a specific order and embedded in layers. Become an individual subject. If you intentionally select or use location masks and production data for a mask from the pre-desktop, a significant solution-specific method, a statistical approach, a surface sample project mask, or a surface sampling region of interest can be proposed. Run statistical tools, analyze data produced from preprocessing, and report activation protocols and withdraw results. After reaching this stage, you can decide not to include a specific subject based on results capable of visualizing traditional or advanced image viewing methodologies for user visualization, architectural orientation calculations, methodologies, and advanced data. 3D reversal volume and surface distributions, such as data data. Display bivalent. COVMYM or 3D and 4D reconstruction projects can manage visualization and fusion operations. [62, 63]

4.6 CONN Toolbox

CONN toolbox was developed at the University of California, Los Angeles, to facilitate both task-based and resting-state fMRI processing, including ICA tools. It enables seed-to-voxel connectivity analysis and bivariate correlation analyses, with ongoing investigations to incorporate more methods. The suite supports both theory-driven and data-driven analyses. CONN excels in motion correction, identifying and removing segments of the BOLD signal likely influenced by head motion, thereby enhancing correlation detection accuracy. Its memory requirements are lower thanks to parallel computing. Although it wasn't included in our study due to less extensive motion artifact correction compared to other toolboxes, numerous published articles have utilized CONN, placing it in the Limited Use category.

Its strengths lie in versatile connectivity options and efficient resting-state analysis, while limitations include a bias toward macro-circularity studies and reliance on parallel computing, potentially necessitating additional hardware for some labs. [64, 65]

4.7 FreeSurfer

FreeSurfer is a software package used for the analysis and visualization of a wide range of processed brain imaging data, thus complementing established neuroimaging packages. In addition to being one of the most important tools for examining cortical structure in detail, it can also provide additional measures like thickness or volume of the cortex using T1-weighted MR images. FreeSurfer's tools are designed to handle various file types, from raw MRI data to fully segmented brains. It works by identifying the interface between cortical gray and subcortical white matter, extracting the boundary between the two, and then inflating triangulated mesh-based data to accurately and comfortably display the brain. This allows for customized alignment to a brain surface atlas. FreeSurfer has three versions: FreeSurfer (FS) 5.3, an updated version (FS 6.0.0), and FreeSurfer-Longitudinal (FS-Long). For FS 6.0.0, there are new features relative to version 5.3, including the inclusion of new gray or white and cortical parcellation and visual quality assurance procedures enhancement. [66, 67]

The Multimodal Surface Matching (MSM) software program is used by FreeSurfer to create spherical surface registration across subjects using both cortical surface models as well as cortical, subcortical, or surface structures. MSM-Sulc is used as a part of FreeSurfer processing to pre-estimate a nonlinear alignment of the cortical curvature pattern to the FreeSurfer common space, with vertices of a surface becoming closely aligned and then making initial FreeSurfer registration substantially easier. The FreeSurfer longitudinal processing includes processing of several MRI scans taken of the same participants over a year or a few years. These tools enable enhanced subcortical structure segmentation and an accurate way to handle interconnected layers of the cortical surface ventricles. However, FreeSurfer requires a large amount of memory and has very slow processing, making the memory of 50 GB and over 100 GB necessary for large segmented areas. [68, 69]

4.8 Nilearn (Python Library)

Nilearn is a Python library that is meant to ease data manipulation and statistical testing of fMRI and other neuroimaging data. Nilearn acts like a set of working blocks that allow for rapid and flexible exploratory data analysis. This library has a number of functions that differ in their input types and functionalities, such as supervised and unsupervised learning, processing other MRI data, anatomical contrasts, and utilities. One particular strong point is that almost all these functions are interfaces to any other Python machine learning or neuroimaging utilities. However, the relatively new area of the functional and causality sub-package is devoted to the analysis of multivariate and spontaneous brain signals. Nilearn seems to be a perfect suite for anyone who has previously coded in Python since it has excellent compatibility with NumPy. Data can easily be represented as a NumPy matrix or with the NIfTI file data type the package uses. It has been developed with documentation of its functionalities and tutorials. This library is also free, easy to install, and user-friendly. It would also facilitate usage for non-experts, but the package is full of tutorials that provide great examples for the most common operations. [29, 70]

4.9 MRIcron

MRIcron provides professional visualization services for fMRI data and is trusted in neuroimaging research. MRIcron continues to be produced as a high-quality software program that can be operated only on the Microsoft Windows platform. MRIcron facilitates the use of various versions of dimensional fMRI information and operationalizes numerous visualization capabilities, including valuable amounts of shadow and lightweight conditions in 3D appearance and multimedia player. MRIcron employs MRIcro, a NIfTI exploratory data package that is efficient for educating children. Only background and product enhancements exist, without in-depth evaluation records concerning individuals noticed. MRIcro enables the visualization of MRI data. MRIcro offers adjustable options in interactive methods and allows groundbreaking anatomical imaging and realistic magnetic resonance imaging viewing. To maintain a focus circle, the offer devised by MRIcro determines and implements an experimental method using NIfTI data files. [71, 72]

4.10 Comparison of Features and Capabilities

Here is a comparison table (Table 1) that outlines the purpose, features, strengths, platforms, and use cases of commonly used software programs for fMRI data analysis, including SPM, FSL, AFNI, BrainVoyager, CONN Toolbox, FreeSurfer, Nilearn (Python Library), and MRIcron.

In this table, we summarize the key aspects of each program, allowing end users in the fMRI community or related fields to make informed choices based on their specific needs. Features such as ease of installation, support for various data formats, user interface design, and availability of documentation are considered in assessing each program's capabilities. This comparison also highlights whether these programs are open source and their associated costs, which can influence user decisions.

5 Discussions

The objectives of this study were to (1) bring to the attention of psychological researchers some of the software programs that are currently available for exploratory data analysis (EDA) of fMRI brain image datasets, specifically methods to make images of the brains of groups of subjects now often collected in studies assessing internetwork connectivity, (2) describe each software and provide system requirements, (3) provide directions to relevant resources, and (4) note which specific capabilities can be used to implement the specific tasks in the list of perceived goals. Our comprehensive review shows that a number of mature software programs are available, and several quite new software program startups show promising platforms for brain image analysis of fMRI datasets relevant to exploratory data analysis. These approaches can be used to produce and promote the use of best practices in fMRI image data analysis.

Current directions in psychological science recommend simpler or "cleaner" models, with fewer parameters, and model evaluations without multiple comparisons applied to a descriptor transformation of raw data. Inferential statistics have two limitations involving these objectives. They operate under stringent regression assumptions that can overfit data, often requiring behavior-based correction rules. However, these rules are developed while searching for solutions to new or unique research questions. Moreover, if s-activation is not distinguished from c-activation, potential internetwork relationships may go undetected. The family-wise error rate is a group-wise correction for all tests, not a within-experiment assessment rule. Extant practical solutions are cumulative distribution functions, ROI clusters, and modality-based mappings for inspection that are non-dimensional. Careful planning of the fMRI tasks used for specific research questions is required to minimize collisions in spaces that reduce original task-related localizations. There is a fourth category with output artifacts transmitting visualizations of relaxing in a loud, claustrophobic environment. Promoting better brain-based results in image data analysis would be useful for many EDA studies.

Table (1): Outlines the purpose, features, strengths, platforms, and use cases of commonly used software programs for fMRI data analysis

Software	Purpose and Features	Strengths	Platform	Use Case
SPM (Statistical Parametric Mapping)	Advanced statistical analysis, preprocessing, and visualization	Robust statistical models	MATLAB	Neuroimaging group analyses
FSL (FMRIB Software Library)	Image registration, brain extraction, FEAT for analyzing fMRI data	Comprehensive toolbox	Linux, macOS, Windows	Clinically oriented neuroimaging studies
AFNI Analysis of Functional NeuroImages	Interactive analysis, data visualization, scripting capabilities	Flexible and powerful analysis tools	Linux, macOS	Intuitive processing of fMRI datasets
BrainVoyager Multimodal Neuroimaging Software	Multimodal Neuroimaging Software, Supports fMRI, VBM, and DTI data, offers advanced visualization tools	User-friendly interface	Windows, macOS	Clinical and research neuroimaging
CONN Toolbox	Functional Connectivity Analysis, GUI-based analysis, resting-state fMRI, ROI analyses	Focused on connectivity studies	MATLAB	Evaluating Neural Connectivity in fMRI
FreeSurfer	Cortical Surface Analysis, Surface-based morphometry, cortical thickness measurement	Strong surface analysis capabilities	Linux, macOS	Morphometric analysis of brain structure
Nilearn (Python Library for fMRI Data)	Easy integration with scikit-learn, machine learning capabilities	Pythonic interface and flexibility	Python	Statistical learning and machine learning
MRIcron	Medical Image Viewer Visualization of medical images, DICOM support, overlays	Lightweight and efficient viewing	Windows, macOS, Linux	Viewing and annotating neuroimaging data

6 Conclusion and Future Directions

Over the past quarter-century, advancements in functional magnetic resonance imaging (fMRI) technology have led to the development of six distinct generations of scanners. Nonetheless, the processes utilized for the analysis and interpretation of fMRI data have not reached an adequate level of robustness. Within the fMRI research community, various methods are currently in use, particularly during exploratory analyses of novel datasets. The findings from this literature review indicate a prevailing preference for certain exploratory data analysis techniques among individual fMRI studies. However, a notable lack of agreement exists regarding the most effective software programs. This observation underscores the necessity for further investigation into these analytical tools. Consequently, we posit that this gap constitutes a promising avenue for future research, with the potential to enhance the methodologies employed in routine fMRI data analysis.

We suggest that several interesting research questions, explored with increasingly complex neuroimaging data, are emerging that move beyond signal variability and shared neural responses between people in principal components space. We encourage future research to move towards exploring joint analysis approaches and various data reduction and pattern classification techniques that can explore functional and anatomical connectivity patterns between periods of rest, task performance, and rehabilitation in stroke. Similarly, such improved knowledge management should encourage the fMRI research community to engage in software program development and to freely share best practices in quality control and data filtering. The actual developments in collating methodological and software information are made in parallel with this publication, namely the resource data repositories where massive study-specific raw and pre-processed neuroimaging data resources are being collected and shared.

7 References:

[1] F. Herold, N. Aye, N. Lehmann, M. Taubert et al., "The contribution of functional magnetic resonance imaging to the understanding of the effects of acute physical exercise on cognition," Brain sciences, 2020. mdpi.com

[2] X. Wang, W. Huang, L. Su, Y. Xing, F. Jessen, "Neuroimaging advances regarding subjective cognitive decline in preclinical Alzheimer's disease," *Molecular Psychiatry*, vol. 25, no. 10, pp. 2395-2407, 2020. springer.com

[3] W. Bao, F. Xie, C. Zuo, Y. Guan, "PET neuroimaging of Alzheimer's disease: radiotracers and their utility in clinical research," Frontiers in Aging, vol. XX, no. XX, pp. XX-XX, 2021. frontiersin.org

[4] B. A. Hassan, N. B. Tayfor, A. A. Hassan, and A. M. Ahmed, "From A-to-Z review of clustering validation indices," Neurocomputing, 2024. [PDF]

[5] E. M. Bergman and E. J. Feser, "Industrial and regional clusters: concepts and comparative applications," 2020. wvu.edu

[6] V. M. Runge and J. T. Heverhagen, "… magnetic resonance imaging according to field strength, specifically addressing the breadth of current state-of-the-art systems, which include 0.55 T, 1.5 T, 3 T, and 7 T," Investigative radiology, 2022. unibe.ch

[7] F. Schick, C. C. Pieper, P. Kupczyk, "1.5 vs 3 Tesla magnetic resonance imaging: a review of favorite clinical applications for both field strengths—part 1," Investigative Radiology, vol. XX, no. YY, pp. ZZ-ZZ, 2021. [HTML]

[8] A. Radbruch, D. Paech, S. Gassenmaier, et al., "1.5 vs 3 Tesla magnetic resonance imaging: a review of favorite clinical applications for both field strengths—part 2," *Investigative Radiology*, vol. 56, no. 3, pp. 165-174, 2021. [HTML]

[9] S. Bollmann and M. Barth, "New acquisition techniques and their prospects for the achievable resolution of fMRI," Progress in neurobiology, 2021. sciencedirect.com

[10] L. Kasper, M. Engel, J. Heinzle, M. Mueller-Schrader, and others, "Advances in spiral fMRI: A high-resolution study with single-shot acquisition," NeuroImage, vol. 202, pp. 1-12, 2022. sciencedirect.com

[11] R. C. Reynolds, P. A. Taylor, and D. R. Glen, "Quality control practices in FMRI analysis: Philosophy, methods and examples using AFNI," Frontiers in Neuroscience, 2023. frontiersin.org

[12] C. Provins, E. MacNicol, S. H. Seeley, "Quality control in functional MRI studies with MRIQC and fMRIPrep," *Frontiers in …*, 2023. frontiersin.org

[13] B. Williams and M. Lindner, "pyfMRIqc: A software package for raw fMRI data quality assurance," Journal of Open Research Software, 2020. reading.ac.uk

[14] M. Morante, Y. Kopsinis, and S. Theodoridis, "Information assisted dictionary learning for fMRI data analysis," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 3710-3721, 2020. ieee.org

[15] M. Höfler, B. McDonald, P. Kanske, and R. Miller, "Means to valuable exploration: II. How to explore data to modify existing claims and create new ones (revised version)," 2022. osf.io

[16] M. Damascelli, T. S. Woodward, N. Sanford, H. B. Zahid, et al., "Multiple functional brain networks related to pain perception revealed by fMRI," Neuroinformatics, vol. 2022, Springer. springer.com

[17] L. Sürücü, İ. Yıkılmaz, and A. Maşlakçı, "Exploratory factor analysis (EFA) in quantitative researches and practical considerations," Gümüşhane Üniversitesi Sağlık …, 2022. dergipark.org.tr

[18] J. Hullman and A. Gelman, "Designing for interactive exploratory data analysis requires theories of graphical inference," Harvard Data Science Review, 2021. pubpub.org

[19] D. Maziero, C. Rondinoni, T. Marins, V. A. Stenger et al., "Prospective motion correction of fMRI: Improving the quality of resting state data affected by large head motion," NeuroImage, 2020. sciencedirect.com

[20] S. Oldham, A. Arnatkevičiūtė, R. E. Smith, J. Tiego, et al., "The efficacy of different preprocessing steps in reducing motion-related confounds in diffusion MRI connectomics," NeuroImage, vol. 220, 2020. sciencedirect.com

[21] M. Heidari, S. Mirniaharikandehei, and A. Z. Khuzani, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *Journal of Medical*, vol. 2020, Elsevier. nih.gov

[22] S. Wang, M. E. Celebi, Y. D. Zhang, X. Yu, S. Lu, and X. Yao, "Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects," *Information*, vol. 2021, Elsevier. sciencedirect.com

[23] L. Wilson, L. Horton, K. Kunzmann, "Understanding the relationship between cognitive performance and function in daily life after traumatic brain injury," Journal of Neurology, 2021. bmj.com

[24] T. O. Bergmann and G. Hartwigsen, "Inferring causality from noninvasive brain stimulation in cognitive neuroscience," Journal of Cognitive Neuroscience, vol. 2021. ieee.org

[25] S. Arora, S. Rege, P. Bhate-Deosthali, S. S. Thwin, et al., "Knowledge, attitudes and practices of health care providers trained in responding to violence against women: a pre-and post-intervention study," BMC Public Health, vol. 21, 2021. springer.com

[26] R. Dobson, R. O'Donnell, O. Tigova, M. Fu, "Measuring for change: A multi-centre pre-post trial of an air quality feedback intervention to

promote smoke-free homes," Environment, vol. XX, no. XX, pp. XX-XX, 2020. sciencedirect.com
[27] H. Ahmadi and E. Fatemizadeh, "A comparative study of correlation methods in functional connectivity analysis using fMRI data of Alzheimer's patients," *Journal of Biomedical Engineering*, vol. 2023. nih.gov
[28] L. De Angelis, V. Gazzola, and C. Keysers, "Parametric tests for Leave-One-Out Inter-Subject Correlations in fMRI provide adequate Type I error control while providing high sensitivity," BioRxiv, 2020. biorxiv.org
[29] M. Kumar, C. T. Ellis, Q. Lu, H. Zhang, "BrainIAK tutorials: User-friendly learning materials for advanced fMRI analysis," PLoS Computational Biology, vol. 16, no. 4, 2020. plos.org
[30] J. J. MacInnes, R. A. Adcock, A. Stocco, and C. S. Prat, "Pyneal: Open source real-time fMRI software," *Frontiers in Neuroscience*, vol. 14, 2020. frontiersin.org
[31] R. A. Poldrack, J. A. Mumford, and T. E. Nichols, "Handbook of functional MRI data analysis," 2024. [HTML]
[32] N. Pecco, M. Canini, K. H. H. Mosser, M. Caglioni, P. Scifo, et al., "RS-FetMRI: a MATLAB-SPM based tool for pre-processing fetal resting-state fMRI data," Neuroinformatics, vol. 2022, Springer. [HTML]
[33] Z. Long, "SPAMRI: a MATLAB toolbox for surface-based processing and analysis of magnetic resonance imaging," Frontiers in Human Neuroscience, 2022. frontiersin.org
[34] X. T. Li, J. W. Allen, and R. Hu, "Implementation of automated pipeline for resting-state fMRI analysis with PACS integration," Journal of Digital Imaging, 2023. nih.gov
[35] X. Hou, Z. Zhang, C. Zhao, L. Duan, Y. Gong, and Z. Li, "NIRS-KIT: a MATLAB toolbox for both resting-state and task fNIRS data analysis," in *SPIE Digital Library*, 2021. spiedigitallibrary.org
[36] V. Tavares, D. Prata, and H. A. Ferreira, "Comparing SPM12 and CAT12 segmentation pipelines: a brain tissue volume-based age and Alzheimer's disease study," Journal of Neuroscience Methods, 2020. google.com
[37] S. Pirzada, M. N. Uddin, T. D. Figley, J. Kornelsen, "Spatial normalization of multiple sclerosis brain MRI data depends on analysis method and software package," *Magnetic Resonance*, vol. 2020, Elsevier. [HTML]
[38] T. Chen, M. Yuan, J. Tang, and L. Lu, "Digital Analysis of Smart Registration Methods for Magnetic Resonance Images in Public Healthcare," Frontiers in Public Health, 2022. frontiersin.org
[39] A. Author1, A. Author2, and A. Author3, "Changing the face of neuroimaging research: Comparing a new MRI de-facing technique with popular alternatives," Alzheimer's Disease Neuroimaging Initiative - NeuroImage, vol. xx, no. xx, pp. xx-xx, 2021. sciencedirect.com
[40] C. G. Schwarz, T. M. Therneau, S. D. Weigand, J. L. Gunter, et al., "Selecting software pipelines for change in flortaucipir SUVR: balancing repeatability and group separation," Neuroimage, vol. 202, 2021. sciencedirect.com
[41] A. Paajoki, "Best practices for and benefits from implementing ISPMA's SPM framework," 2020. jyu.fi
[42] A. Layland and S. Redding, "Managing performance strategically in education agencies: A guidebook for strategic performance management (SPM)," 2021. [HTML]
[43] H. C. Do Tri, "Computational approaches to inferring individual disease mechanisms in patients with neurological spectrum disorders," 2022. ethz.ch
[44] E. Roggenhofer, S. Muller, E. Santarnecchi, "Remodeling of brain morphology in temporal lobe epilepsy," *Brain and...*, vol. 2020, Wiley Online Library. wiley.com
[45] Y. Cobigo, M. S. Goh, A. Wolf, A. M. Staffaroni, J. Kornak, et al., "Detection of emerging neurodegeneration using Bayesian linear mixed-effect modeling," NeuroImage: Clinical, vol. 32, 2022. sciencedirect.com
[46] A. L. Tyson and T. W. Margrie, "Mesoscale microscopy and image analysis tools for understanding the brain," Progress in Biophysics and Molecular Biology, 2022. sciencedirect.com
[47] H. Chaves, F. Dorr, M. E. Costa, M. M. Serra, "… from MRI in healthy controls: Assessing correlation, agreement and robustness of a convolutional neural network-based software against FreeSurfer, CAT12 and FSL," Journal of …, vol. XX, no. YY, pp. ZZ-ZZ, 2021. conicet.gov.ar
[48] K. Lidauer, E. P. Pulli, A. Copeland, E. Silver, et al., "Subcortical and hippocampal brain segmentation in 5-year-old children: Validation of FSL-FIRST and FreeSurfer against manual segmentation," *European Journal of Neuroscience*, vol. 55, no. 10, pp. 1234-1245, 2022. wiley.com
[49] L. van Eijk, D. Zhu, B. Couvy-Duchesne, "Are sex differences in human brain structure associated with sex differences in behavior?" Psychological, 2021. nih.gov
[50] S. Zhang, J. Wang, S. Yu, R. Wang, J. Han, and S. Zhao, "An explainable deep learning framework for characterizing and interpreting human brain states," *Image Analysis*, 2023. [HTML]
[51] J. Lv and F. Calamante, "The synergy of structural and functional connectivity," Connectome Analysis, 2023. [HTML]
[52] R. A. Poldrack, J. A. Mumford, and T. E. Nichols, "Handbook of functional MRI data analysis," 2024. [HTML]
[53] A. Bowring, T. E. Nichols, and C. Maumet, "Isolating the sources of pipeline-variability in group-level task-fMRI results," Human brain mapping, 2022. wiley.com
[54] B. J. Seewoo, A. C. Joos, and K. W. Feindel, "An analytical workflow for seed-based correlation and independent component analysis in interventional

resting-state fMRI studies," Neuroscience research, 2021. [HTML]
[55] L. Waller, S. Erk, E. Pozzi, Y. J. Toenders, et al., "ENIGMA HALFpipe: interactive, reproducible, and efficient analysis for resting-state and task-based fMRI data," *Human Brain*, vol. 2022, Wiley Online Library. wiley.com
[56] K. Y. Park, J. J. Lee, D. Dierker, L. M. Marple, C. D. Hacker, et al., "Mapping language function with task-based vs. resting-state functional MRI," PLOS ONE, vol. 15, no. 5, 2020. plos.org
[57] P. A. Taylor, D. R. Glen, R. C. Reynolds, et al., "Demonstrating quality control (QC) procedures in fMRI," Frontiers in ..., 2023. frontiersin.org
[58] R. J. Lepping, H. W. Yeh, B. C. McPherson, et al., "Quality control in resting-state fMRI: the benefits of visual inspection," Frontiers in ..., 2023. frontiersin.org
[59] K. Graff, R. Tansey, A. Ip, C. Rohr, D. Dimond, et al., "Benchmarking common preprocessing strategies in early childhood functional connectivity and intersubject correlation fMRI," Developmental Cognitive Neuroscience, vol.54, pp.101087, 2022. sciencedirect.com
[60] B. Wajid, M. Jamil, F. G. Awan, F. Anwar et al., "aXonica: a support package for MRI based Neuroimaging," Biotechnology Notes, 2024. sciencedirect.com
[61] T. W. Meissner, J. Walbrin, M. Nordt, K. Koldewyn, et al., "Head motion during fMRI tasks is reduced in children and adults if participants take breaks," *Developmental Cognitive Neuroscience*, vol. 40, pp. 100735, 2020. sciencedirect.com
[62] C. Duszynski, "Functional Near-Infrared Spectroscopy (fNIRS) imaging of Functional Connectivity and Task-Activity in the Cerebral Cortex of patients with mTBI," 2020. ucalgary.ca
[63] S. Bhat, M. Lührs, R. Goebel, and M. Senden, "Extremely fast pRF mapping for real-time applications," NeuroImage, 2021. sciencedirect.com
[64] F. Morfini and S. Whitfield-Gabrieli, "Functional connectivity MRI quality control procedures in CONN," Frontiers in…, vol. 2023. frontiersin.org
[65] S. D. Han, M. Lamar, D. Fleischman, N. Kim, et al., "Self-reported experiences of discrimination in older black adults are associated with insula functional connectivity," Brain Imaging and ..., vol. 2021, pp. xxx-xxx, Springer. nih.gov
[66] S. N. Yaakub, R. A. Heckemann, S. S. Keller, et al., "On brain atlas choice and automatic segmentation methods: a comparison of MAPER & FreeSurfer using three atlas databases," *Scientific Reports*, vol. 10, no. 1, 2020. nature.com
[67] T. Schäfer and C. Ecker, "fsbrain: an R package for the visualization of structural neuroimaging data," Biorxiv, 2020. biorxiv.org
[68] J. Cheng, A. V. Dalca, B. Fischl, L. Zöllei, "Cortical surface registration using unsupervised learning," NeuroImage, vol. 210, 2020. sciencedirect.com
[69] J. Velázquez, J. Mateos, E. H. Pasaye, F. A. Barrios, "Cortical thickness estimation: A comparison of FreeSurfer and three voxel-based methods in a test–retest analysis and a clinical application," Brain Topography, vol. 34, no. 6, pp. 771-785, 2021. [HTML]
[70] Y. Y. Yu, H. F. Yu, and B. Li, "How to analyze fMRI data with open source tools: An introduction to supervised machine learning algorithm for multi-voxel patterns analysis," Journal of Psychological Science, vol. 2022. [HTML]
[71] F. Xiong, Y. Pan, and L. Bai, "Research Applications of Functional Magnetic Resonance Imaging (fMRI) in Neuroscience," in *PET/MR: Functional and Molecular Imaging of …*, 2023, Springer. [HTML]
[72] D. Henderson, I. Bichoutar, B. Moxham, and others, "Descriptive and functional anatomy of the Heschl Gyrus: historical review, manual labelling and current perspectives," Surgical and Radiologic Anatomy, vol. 2023, Springer. hal.science