

Multi-voxel pattern analysis of human brain fMRI data

Leesa Joyce

MS14108

*A dissertation submitted for the partial fulfillment of
BS-MS dual degree in Science*



Indian Institute of Science Education and Research Mohali

April 2019

Certificate of Examination

This is to certify that the dissertation titled **Multi-voxel pattern analysis of human brain fMRI data** submitted by **Ms. Leesa Joyce (Reg. No. MS14108)** for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Prof. Somdatta Sinha
(Supervisor)

Dr Arpan Banerjee
(Co-Supervisor)

Dr. Samarjit Bhattacharya

Dr. Shashi Bhushan Pandit

Dated: April 24, 2019

Declaration

The work presented in this dissertation has been carried out by me with Prof.Somdatta Sinha at the Indian Institute of Science Education and Research Mohali and Dr.Arpan Banerjee at the National Brain Research Centre. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Leesa Joyce
(Candidate)

Dated: April 24, 2019

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Prof. Somdatta Sinha
(Supervisor)

Dr. Arpan Banerjee
(Co-supervisor)

Acknowledgements

Firstly, I would like to thank my thesis supervisor Prof. Somdatta Sinha for her guidance and encouragement throughout the duration of my project. The door to Prof. Sinhas office was always open for me to clarify doubts and get suggestions. She motivated me to inculcate an organized way of doing research and taught me the first lessons which are imperative in todays research.

I would also like to thank Dr. Arpan Banerjee and Dr. Dipanjan Ray of the cognitive Brain Dynamics lab at the National Brain Research Centre, Manesar, for hosting me for an entire semester during my dissertation. I am gratefully indebted to their support, guidance and very valuable comments on this thesis.

I must express my very profound gratitude to Puneeth Deraje and Himanshu Agarwal for their extensive support throughout the project. I would like to thank Anagh Pathak for all the helpful discussions, Nisha Sasthry and Vivek Sharma and my lab mates for their significant contributions. Without their passionate participation and input, the project could not have been successfully completed.

My life at IISER Mohali would not have been this fruitful, if not for my great friends (*Othlas*, batch-mates and others), who have always been there to motivate me and so I will always be indebted to them. I would also like to thank my family for their constant support and the unconditional trust they have in me. Finally, I thank IISER Mohali and NBRC Manesar for all the facilities and DST-INSPIRE for the financial support. Thank you.

Leesa Joyce

Contents

Certificate of Examination	iii
Declaration	v
Acknowledgements	vii
List of Figures	xi
Abstract	xv
1 Introduction	1
1.1 History	1
1.2 fMRI and the BOLD signal	2
1.3 Visual Perception	4
1.4 What is Recognition?	5
1.4.1 Object Recognition	5
1.4.2 Face Recognition	5
2 Data-set and Methods	7
2.1 Data and its Origin	7
2.1.1 Participants	7
2.1.2 Experimental Design, Stimuli, and Tasks	7
2.1.3 fMRI Data Acquisition	8
2.1.4 Practice Sessions	8
2.1.5 Pre-processing	9
2.2 The Two Types of fMRI Data Analysis	9
2.2.1 Classical univariate analysis	9
2.2.2 Multi-Voxel Pattern Analysis (MVPA)	10
2.3 MVPA	10
2.3.1 Steps to perform MVPA	11
2.4 Classifiers	11

2.4.1	Support vector machine (SVM)	11
2.4.2	Linear discriminant analysis (LDA)	16
2.4.3	Classification accuracy	18
2.5	Methods of Analysis	19
2.5.1	Region of Interest (ROI) based MVPA	19
2.5.1.1	Statistical significance test	20
2.5.2	Searchlight MVPA	20
2.5.2.1	Statistical analysis	21
2.6	Software and Tools	21
2.6.1	CoSMoMVPA	21
2.6.2	Statistical Parametric Mapping (SPM)	21
2.6.3	MRICRON	22
3	Results	23
3.1	ROI Analyses	23
3.2	Searchlight Analyses	25
4	Discussion and Future Direction	31
4.1	Possible Explanation for the Results	31
4.1.1	ROI Analysis	31
4.2	Interpreting classification accuracy	31
4.2.1	Underestimation	32
4.2.2	Overestimation	32
4.2.3	Other limitations	33
4.3	Conclusions	33
4.4	What next?	33

List of Figures

1.1	Theoretical BOLD signal response. Image taken from [5]	2
1.2	fMRI image of an activated region of the brain	2
1.3	fMRI time series of a voxel, say V. Each of the bigger cubes (1,2,...,T) represent one scan of the brain and in each scan the shaded voxel is V. The image is taken from [6]	3
1.4	Primary visual pathway [7]	4
1.5	Visual dual streams [8]	4
1.6	The visual areas involved in (face and colour) recognition [13]	6
2.1	An example of MVPA for 2 classes [17]	10
2.2	Classification in MVPA [17]	10
2.3	SVM on 2-D dataset with 2 classes	12
2.4	LDA on 2-D data-set with 2 classes [19]	16
2.5	Classification accuracy.	19
3.1	Regions of interest in the brain. We have a combined mask for primary and secondary visual regions (V1)	24
3.2	Significance test for the SVM classification on colour perception.	24
3.3	Significance test for the LDA classification on colour perception.	24
3.4	Significance test for the SVM classification on face perception.	25
3.5	Significance test for the LDA classification on face perception.	25
3.6	Color Perception	27
3.7	Face Perception	29

Dedicated to Prof. Somdatta Sinha

Abstract

This work is aimed at exploring multi-variate pattern analysis of functional magnetic resonance imaging data of human brain as an alternative for situations/questions where the classical univariate analysis falls short. Pattern-based fMRI analysis enables us to address content-based processing in human brain and helps us study the direct link between multivoxel fMRI activity patterns and the corresponding cognitive representations. The first goal is to categorise the mental representations for colour and face perceptions, based on the pattern vectors. Region-of-interest based multivariate pattern analysis shows that there are distinct mental representations for colour and face perceptions, in the primary and secondary visual cortex and the ventral stream. These representations are classified into distinct categories by supervised machine learning algorithms such as support vector machine and linear discriminant analysis. The second part of the work is aimed at a whole brain searchlight analysis to validate the ROI based results. Searchlight analysis enabled us to look at the spatial encoding of task relevant information across the whole brain. In addition to this, efforts have been made to look at transformation of these mental representations with practice.

Chapter 1

Introduction

In simple words cognition refers to thinking. Playing chess, writing a poem, solving mathematical problems, etc are the obvious applications of conscious reasoning-but thought assumes many other subtler forms, such as processing and interpreting sensory input, guiding motor functions, and showing emotions. Research on cognition addresses questions on attention, logical reasoning, memory formation and storage, learning of language, acquisition and retention of knowledge and motor control and many more [1].

This field of research combines the experiments in cognitive psychology with neuroimaging techniques to understand the functioning of the brain and the corresponding cognitive activities. Neuroimaging technology can be used for functional brain imaging to measure an aspect of the function of brain, often with an aim to understanding the link between activity in certain regions of brain and specific cognitive functions [2].

1.1 History

During the 1880s, the first non-invasive neuroimaging technology - human circulation balance - was invented by Angelo Mosso to measure to measure the redistribution of blood during intellectual and emotional activity [3]. There were series of inventions

and improvisations after this, which led to the development of magnetic resonance imaging (MRI) and computed tomography (CT) in the 1970s and 1980s. Next came functional MRI (fMRI), which allowed us to create images which are not just static but can map the brain function to observe the cognitive activities. Due to lack of exposure to radiation, non-invasiveness and relatively wide availability, fMRI has evolved to dominate the neuroimaging field since 1990s [4].

1.2 fMRI and the BOLD signal

We know that when an area of the brain is in use, the neurons in that region are active. Neurons do not have internal energy reserves in the form of sugar and oxygen. So when they fire, they need more energy to be imported quickly. Blood releases excess oxygen to those firing neurons through a process called the hemodynamic response. As a consequence, oxygen is depleted from the surrounding blood vessels. Immediately, the body overcompensates for the oxygen depletion by increasing the flow of oxygenated blood to that region (Figure 1.1).

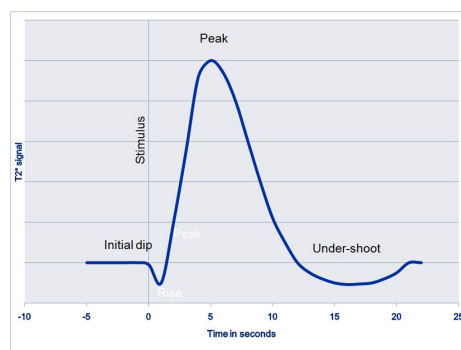
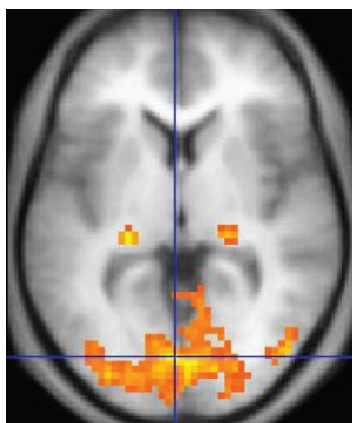


FIGURE 1.1: Theoretical BOLD signal response. Image taken from [5]



The color scale from red to yellow represents activity from low to high against the basal level activity represented by the grey scale. The image is taken from [5]

FIGURE 1.2: fMRI image of an activated region of the brain

Functional magnetic resonance imaging (fMRI) is a technique which measures the activity of brain using these changes associated with cerebral blood flow as a proxy to neuronal activity. The functional image of the brain (Figure 1.2) shows which parts of the brain are involved in a particular cognitive process.

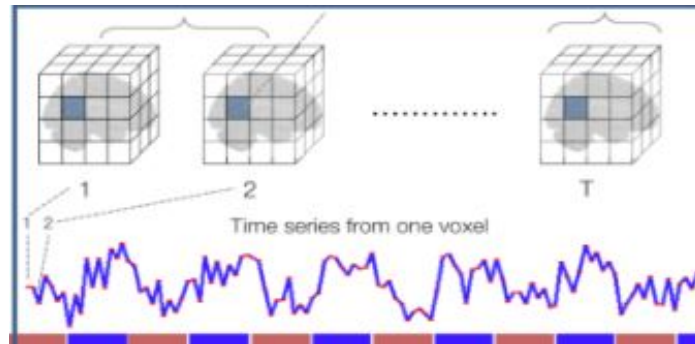


FIGURE 1.3: fMRI time series of a voxel, say V .

Each of the bigger cubes (1,2,...,T) represent one scan of the brain and in each scan the shaded voxel is V . The image is taken from [6]

The signal detected by the scanner is known as the Blood Oxygen-Level Dependent (BOLD) signal. Hemoglobin in its two forms (oxygenated and deoxygenated), has different magnetic properties. Oxy-hemoglobin is diamagnetic and deoxy-hemoglobin is paramagnetic. An MRI scanner can detect the magnetic signal variation due to this inherent difference in their magnetic properties. This difference is the origin of the BOLD signal per voxel which can be captured across time points (Figure 1.3).

A voxel is one of many from which a volume is composed. Technically, voxels are discrete elements into which a representation of a three-dimensional space is divided.

In a typical fMRI experiment, the task will have many repetitions of a thought or action. So, we can use statistical methods to identify the regions of the brain which reliably are most active during that thought or action. Hence, the data-set is a 4 dimensional construct involving 3D brain volumes at different time points. Time being the 4th dimension, we can extract the time-series for each voxel for analysis.

1.3 Visual Perception

It is well known that after primary visual areas (Figure 1.4), visual / visuo-motor information are processed along two distinct neural pathways. The existence of two distinct neural streams (Figure 1.5) - ventral and dorsal, projecting from the primary visual areas to the inferior temporal cortex and the posterior parietal cortex respectively, is known for almost 30 years. Visual dual stream has also motivated discovery of similar duplex architecture in brain regions associated with other cognitive domains like auditory, haptic, and olfactory perception, language, and attention. Thus, it has constituted a general framework for understanding the functional organization of cerebral cortex [4].

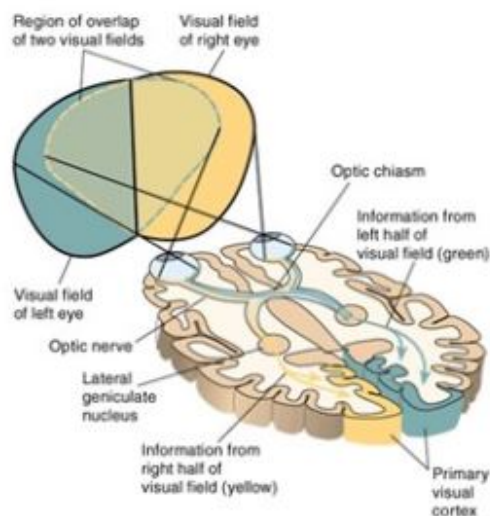


FIGURE 1.4: Primary visual pathway [7]

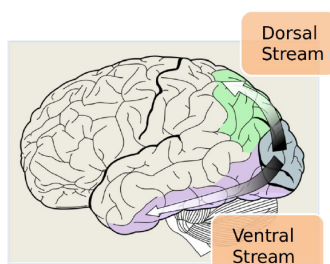


FIGURE 1.5: Visual dual streams [8]

Among the two most influential models of visual dual stream, the MU model proposed by Mishkin and Ungerleider (1982) suggests that the input information decides the neural pathway for processing. Features that help in object identification ("*what*") like color, shape, texture etc. are processed in the ventral stream whereas spatial ("*where*") information (e.g., position, velocity, depth, orientation) take the dorsal stream [9].

In contrast, the Milner-Goodale (MG) model suggests that the output or the task goal decides the processing pathway [10]. The ventral stream areas are needed for internal representation (perception) of both what and where information whereas the dorsal stream is recruited for processing those same input information for guiding action online ("*action*"). However, it is true that in many everyday situations both

models predict same brain activation. For example, ventral stream activation during "perception" of "what" information [9].

1.4 What is Recognition?

In simple words, recognition is a tally between the visual stimuli (processed through the ventral stream) and a mental representation of a physical entity. Though not trivial, luckily we can recognize a visual input from several viewpoints, even when the images of the physical entity are entirely different [11].

1.4.1 Object Recognition

Object recognition is a property of an organism to comprehend the physical properties (shape, texture, and color) of an object and assign semantic attributes to it (for example, identifying the object as an apple). This ability includes, understanding the use of the object, its previous encounter, and how it relates to other people in case of humans. Irrespective of an object's position and illumination, humans have gained the potential to effectively identify an object and label it. We are one among the very few species to possess the capability of invariant visual object recognition. Research suggests that Lateral Occipital Cortex (LOC) is involved in object recognition. Two types of processing, Top-Down, which is knowledge/goal based, and Bottom-Up, which is sensory driven are required for a species to recognize objects at varying distances, angles, lighting, etc. [12].

1.4.2 Face Recognition

There have been strong evidences to show that visual cortex in the brain has discrete areas involved in processing faces unlike for any other objects. It has also been reported that a separate neural circuitry is required for identifying individual faces in a viewpoint independent manner. There are separate regions in the brain to process

faces and objects which are spatially separated by several centimeters. The region

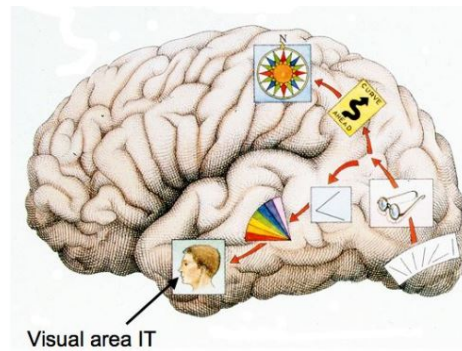


FIGURE 1.6: The visual areas involved in (face and colour) recognition [13]

of the brain responsible for processing faces is known as the infero-temporal (IT) cortex (Figure 1.6), located in the ventral lobe. Specifically, *fusiform gyrus* is the one responsible for processing color information and for face-body recognition [14].

Chapter 2

Data-set and Methods

2.1 Data and its Origin

The data used in this project is taken from the Cognitive Brain Dynamics lab at the National Brain Research Centre, Manesar. (<https://cognitivebrainlab.weebly.com>)

2.1.1 Participants

The data is obtained from 20 young, healthy subjects (Mean age = 25.35 years, SD=2.796 years, 13 females, 7 males). Participants were all right-handed according to the Edinburg Inventory with normal or corrected-to-normal vision. They gave their written informed consent to the experimental procedure, which was approved by the Ethics Committee of National Brain Research Center (NBRC). After preliminary tests, 2 subjects were identified as outliers. Hence, all the analyses are restricted to the other 18 subjects.

2.1.2 Experimental Design, Stimuli, and Tasks

In perception task involving color stimuli, four different color dots were presented sequentially and the participant was asked, at the end of the run, to denote verbally

the number of times the target color (red) were presented. Similarly in face perception task, four different faces were presented and the task was to indicate the number of times a particular target face (that had been shown before the fMRI scan started) was presented. Both the stimuli were presented on the center of the screen. Stimuli were presented in on blocks (duration 24 seconds each) alternating with off blocks of 16 seconds duration. During off blocks a central cross in a grey background was presented. Stimulus order was randomized within an on block. In perception tasks, each stimuli was presented for 2 seconds. In our case, we have 8 runs per task each run containing 20 samples. Among the 20 samples, the first 12 samples are the brain volumes acquired during active blocks (colour or face is shown). The last 8 samples are the brain volumes acquired during the rest block (when the subject is resting).

2.1.3 fMRI Data Acquisition

Images were acquired on a 3T (Philips achieva) magnetic resonance imaging (MRI) scanner at the National Neuroimaging Facility of NBRC (<https://cognitivebrainlab.weebly.com>). The experimental design and the data acquisition is done by Dr. Dipanjan Ray for a project on visual dual streams. A part of that data is used for MVPA in this project.

2.1.4 Practice Sessions

To assess the effect of practice on brain activation, seven practice sessions were conducted across seven separate days between two fMRI scan sessions. Each practice session comprised of same two tasks as scanning sessions and took place in an isolated room in the lab. Stimuli were presented in a computer screen. The order of presentation of stimuli in a task was randomized and was different for different sessions. Overall, the order of two tasks was also randomized. In contrast to the fMRI sessions, tasks were performed in sitting posture, but the distance between the participant and the screen remains same. The number of practice sessions was decided based on a pilot study probing the improvement of response time with practice.

2.1.5 Pre-processing

The pre-processing and statistical analysis of fMRI data were executed with SPM8 (Statistical Parametric Mapping, <http://www.fil.ion.ucl.ac.uk/spm/>) program. Initial 8 seconds of scan was discarded to allow the magnetization to stabilize to a steady state. Prior to statistical analysis, images were slice time corrected, realigned with the mean image, motion corrected, co-registered with the corresponding T1-weighted images, normalized to a Montreal Neurological Institute (MNI) reference template [15] and re-sampled to 4-4-5 mm^3 . Temporal high pass filtering with cut off of 128 seconds was employed to remove low frequency drifts caused by physiological and physical (scanner related) noises.

2.2 The Two Types of fMRI Data Analysis

2.2.1 Classical univariate analysis

fMRI was initially used to study the activity of brain at the level of macro-anatomical regions based on group data. These group data were smoothed spatially and warped anatomically to standard brain templates. The smoothing across voxels and the spatial spread of hemodynamic response leads to local spatial dependencies.

In a classical mass univariate analysis, statistical analysis is performed for these smoothed group data using a general linear model separately for each voxel in the brain. Analysis is done by fitting a model independently to the time series of individual voxel or for the mean signal time series of a Region-Of-Interest (ROI). These smoothed group results are of no use when we are interested to know how individual cognitive representations are encoded and transformed in the human brain [16]. In a mass univariate analysis activation of single voxels is related to the psychological dimensions. So, this classical method cannot map the neural basis of experimental conditions for which the effect on activation is multi-dimensionally distributed [17].

2.2.2 Multi-Voxel Pattern Analysis (MVPA)

Pattern-based fMRI analyses enable us to address content-based processing in the human brain by correlating multi-voxel fMRI activity patterns with the corresponding cognitive representations [17].

2.3 MVPA

fMRI signals provide some sort of representation of the neural signals [18]. However the exact relation between fMRI signal and the neural signal are not established. But what is well-established, is the correlation between the spatial pattern of fMRI scan and the mental state of humans. In order to access this infor-

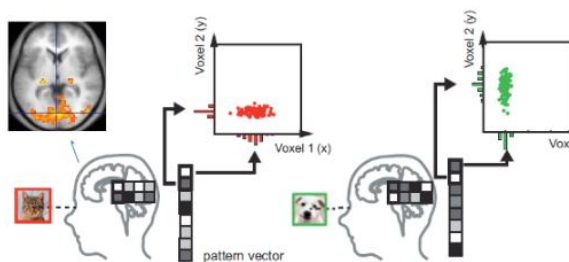


FIGURE 2.1: An example of MVPA for 2 classes [17]

mation, scientists have long used classifiers and explored brain processes at the level of mental representation. Due to vastness and complexity of the brain structure, one needs to be cautious while making interpretations of the fMRI data. This is achieved through pattern based analyses and computational modeling.

During an activity, at any given instance the brain can be thought of as a vector of its voxel values. Hence at any instance it is nothing but a single point in the n -dimensional voxel space. It is now possible to tag these data points with the condition in which they were acquired. Multiple measurements will provide us with a set of points for each condition (Figure 2.1). The aim of the analyses is to be able to classify new data points based on the previously available data. This is achieved through classifiers, which attempt to separate points without over-fitting (Figure 2.2). This then allows one to associate mental representations with the condition.

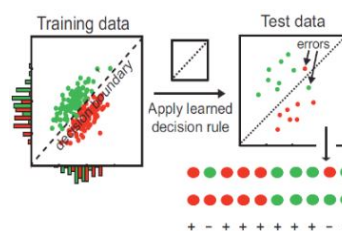


FIGURE 2.2: Classification in MVPA [17]

2.3.1 Steps to perform MVPA

1. **Feature selection:** Restrict your Analysis to voxels which are relevant, based on previous literature and anatomical knowledge of the brain or based on your region of interest.
2. **Pattern assembly:** Rearrange the data in an orderly fashion depending on the sequence of events during the experiment, in a single scan. A single row which has the activity in all the required voxels at a given time is called a pattern.
3. **Classifier Training :** Run the Multivariate classification algorithm on these patterns.
4. **Generalized Testing :** Run your trained algorithm on new data to confirm efficiency of classification.

2.4 Classifiers

Classifiers group the training data into discrete categories, each with a unique label (experimental condition). In this study 2 supervised machine learning algorithms namely Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) are used. For classification, both LDA and SVM use a weight at each voxel to linearly project the data points to a single decision axis. They use different algorithms for estimating the weights from the training data.

2.4.1 Support vector machine (SVM)

Support Vector Machine is a classification technique in which one tries to determine a hyperplane that separates the classes. The hyperplane is chosen so as to maximise the distance from the classes.

Linear Support Vector Machine

Linear Support Vector Machine that is used to for classification when there are precisely two classes and we can find a hyper plane that separates the two classes. For instance for a 2-Dimensional data points this would imply that there exists a line that can separate the two classes completely.

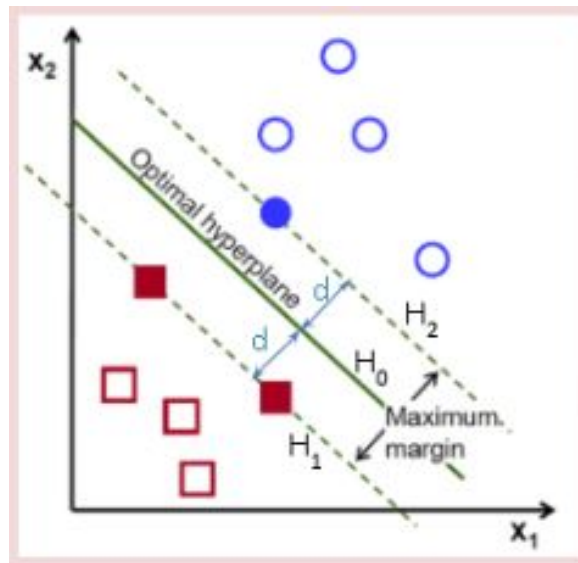


FIGURE 2.3: SVM on 2-D dataset with 2 classes

Method

Assume the separating Hyperplane is given by

$$H_0 : \vec{w} \cdot \vec{x} + b = 0 \text{ where,} \quad (2.1)$$

\vec{w} is the vector perpendicular to H_0 and

$\frac{b}{\|\vec{w}\|}$ is the perpendicular distance of H_0 from the origin

Notice that for 2-D data points the above equation simply represents the equation of a line.

Let H_1 and H_2 be the hyperplane which contain the points of Class 1 and Class 2 respectively that are at the shortest distance from H_0 . Such points are called support vectors. In Figure 2.3 the solid red and blue data points are the support vectors. We

define the set of indices of support vectors as

$$S := \{i \in \{1, 2, 3, \dots, N\} | x_i \in H_1 \text{ or } x_i \in H_2\}$$

Since our aim is to find H_0 such that it is equidistant from H_1 and H_2 , one can vary b and \vec{w} in order to make

$$H_1 : \vec{w} \cdot \vec{x} + b = -1 \quad (2.2)$$

$$H_2 : \vec{w} \cdot \vec{x} + b = +1 \quad (2.3)$$

Let \vec{x}_i be the i^{th} sample data point. To each x_i assign y_i such that

$$y_i = \begin{cases} -1 & \text{if } x_i \text{ belongs to class 1} \\ 1 & \text{if } x_i \text{ belongs to class 2} \end{cases}$$

Therefore our aim is to find \vec{w} and b such that :

1. (2.1) (2.2) and (2.3) holds
2. The distance d between H_0 and H_1 (or H_2) is maximized
3. For every i :

$$\vec{w} \cdot \vec{x}_i + b \leq -1 \text{ if } y_i = -1$$

$$\vec{w} \cdot \vec{x}_i + b \geq 1 \text{ if } y_i = 1$$

Equivalently the above two conditions can be clubbed into one.

$$y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 \quad \forall i$$

Now, Let a be a scalar such that :

$$a \cdot \vec{w} \in H_1$$

$$\begin{aligned}
&\Rightarrow \left(a + \frac{d}{\|\vec{w}\|}\right) \cdot \vec{w} \in H_0 \\
&\Rightarrow (a \cdot \vec{w}) \cdot \vec{w} + b = -1 \quad \& \quad \left(a + \frac{d}{\|\vec{w}\|}\right) \cdot \vec{w} \cdot \vec{w} + b = 0 \\
&\Rightarrow a\|\vec{w}\|^2 + b = -1 \quad \& \quad \left(a + \frac{d}{\|\vec{w}\|}\right)\|\vec{w}\|^2 + b = 0 \\
&\Rightarrow \frac{d}{\|\vec{w}\|}\|\vec{w}\|^2 = 1 \Rightarrow d = \frac{1}{\|\vec{w}\|}
\end{aligned}$$

Therefore our problem is reduced to the following :

Find \vec{w} and b that

$$\text{maximise } \frac{1}{\|\vec{w}\|} \text{ with the constraints } y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 \quad \forall i$$

Equivalently,

Find \vec{w} and b that

$$\text{minimize } \frac{1}{2}\|\vec{w}\|^2 \text{ with the constraints } y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 \quad \forall i$$

To solve this problem we use the Lagrange Multiplier Method and introduce Lagrange Multipliers α_i

$$L_p = \frac{1}{2}\|\vec{w}\|^2 - \sum_{i=1}^N \alpha_i (y_i(\vec{w} \cdot \vec{x}_i + b) - 1) \quad (2.4)$$

Now we need to find b and \vec{w} such that L_p is maximised and α_i such that $\alpha_i \geq 0$ and L_p is minimized. In order to find such b and \vec{w} ,

$$\begin{aligned}
&\frac{\partial L_p}{\partial b} = 0 \quad \& \quad \frac{\partial L_p}{\partial \vec{w}} = 0 \\
&\therefore L_p = \frac{1}{2}\vec{w} \cdot \vec{w} - \sum_{i=1}^N \alpha_i (y_i(\vec{w} \cdot \vec{x}_i + b) - 1) \\
&\therefore \frac{\partial L_p}{\partial b} = \vec{w} - \sum_{i=1}^N \alpha_i y_i \vec{x}_i \quad \& \quad \frac{\partial L_p}{\partial \vec{w}} = - \sum_{i=1}^N \alpha_i y_i b \\
&\therefore \vec{w} = \sum_{i=1}^N \alpha_i y_i \vec{x}_i \quad \& \quad \sum_{i=1}^N \alpha_i y_i b = 0
\end{aligned} \quad (2.5)$$

Substituting this back into (2.4), we get

$$L_p = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (\vec{x}_i \cdot \vec{x}_j)$$

Define $\mathbf{H} := [H_{ij}]_{N \times N}$, where $H_{ij} = y_i y_j (\vec{x}_i \cdot \vec{x}_j)$ & $\boldsymbol{\alpha} = [\alpha_1 \alpha_2 \dots \alpha_N]$

$$\text{Then, } L_p = \sum_{i=1}^N \alpha_i - \frac{1}{2} \boldsymbol{\alpha} \cdot \mathbf{H} \cdot \boldsymbol{\alpha}^T$$

The above term is called the Dual and denoted by L_D

Therefore our problem reduces to finding $\alpha_i \geq 0$ to maximize the Dual L_D . This problem can be solved using the Quadratic Problem Solver Method.

It can be shown that $\alpha_i = 0 \forall i \in S$ i.e for all the support vectors. Intuitively notice that the constraint that $y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1$ is not significant for $i \notin S$. This is because if the data points are below H_1 or above H_2 then if $y_i(\vec{w} \cdot \vec{x}_i + b) = 1 \forall i \in S$ then the inequality will be true for all i , since we have assumed that the data is linearly separable.

Once, the α_i are determined, By (2.5) we have :

$$\vec{w} = \sum_{i=1}^N \alpha_i y_i \vec{x}_i = \sum_{i \in S} \alpha_i y_i \vec{x}_i$$

In order to determine b look at the sample points that lie on the Support Vectors say, x_s , Then :

$$\begin{aligned} y_s(\vec{w} \cdot \vec{x}_s + b) &= 1 \\ \Rightarrow y_s^2(\vec{w} \cdot \vec{x}_s + b) &= y_s \\ \text{But, } y_s^2 &= 1 \\ \therefore b &= y_s - \vec{w} \cdot \vec{x}_s = y_s - \sum_{i \in S} \alpha_i y_i \vec{x}_i \cdot \vec{x}_s \end{aligned}$$

Since $\forall s \in S$ we would get a different value of b , the ideal choice of b would be the average.

$$\therefore b = \frac{1}{|S|} \sum_{s \in S} (y_s - \sum_{i \in S} \alpha_i y_i \vec{x}_i \cdot \vec{x}_s)$$

Thus, we have found the optimal Separating Hyperplane.

$$H_0 : \sum_{i \in S} \alpha_i y_i \vec{x}_i \cdot \vec{x} + \frac{1}{|S|} \sum_{s \in S} (y_s - \sum_{i \in S} \alpha_i y_i \vec{x}_i \cdot \vec{x}_s) = 0$$

2.4.2 Linear discriminant analysis (LDA)

Linear discriminant analysis creates a classification boundary by projecting sample data points in an n -dimensional space on to an appropriate lower dimensional object and then specifying hyper-planes in that lower dimension, that can distinguish between the classes.

For example, if your data points are in a 2 dimensional space and, say, there are 2 classes (Figure 2.4). Then, LDA determines the optimum line (the lower dimensional object) such that when we project our data points onto that line, the two classes can be easily distinguished. Ideally, such a line should be such that the projected points have a low within-class variance and a high between-class variance.

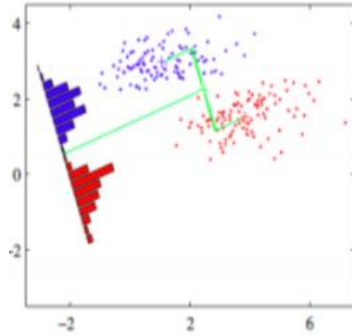


FIGURE 2.4: LDA on 2-D data-set with 2 classes [19]

Method

Suppose the data points are in an n -dimensional space and there are C classes.

Let N_i be the number of data points of class i .

Let $x_{i,j}$ be the j^{th} sample of the i^{th} class, where $1 \leq i \leq C$, $1 \leq j \leq N_i$. Note that $x_{i,j}$ would be an n -dimensional vector.

In order to compute the within-class variance and between class variance, we need the following quantities.

The sample mean of class i .

$$m_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{i,j}$$

The sample mean of the entire samples

$$m = \frac{1}{N} \sum_{i=1}^C N_i m_i = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{N_i} x_{i,j}$$

The Co-variance matrix of Class i

$$Cov_i = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (x_{i,j} - m_i)(x_{i,j} - m_i)^T$$

Notice the similarity between the co-variance matrix and the variance for 1-D data points. In fact the co-variance matrix for 1-D data-set is $Cov_i = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (x_{i,j} - m_i)^2$ which is precisely the variance. Hence, the co-variance matrix of class i gives a measure of the variability within the class i . The within-class scatter S_w , a measure of the variability within classes of the sample data points is given by

$$S_w = \sum_{i=1}^C (N_i - 1) Cov_i = \sum_{i=1}^C \sum_{j=1}^{N_i} (x_{i,j} - m_i)(x_{i,j} - m_i)^T$$

Similarly, the between-class scatter, S_b , a measure of variability between points in different classes is given by

$$S_b = \sum_{i=1}^C (m_i - m)(m_i - m)^T$$

A projection matrix ϕ is matrix which when applied to the points will project it into the required space. Our aim is to find a projection that minimizes the within-class scatter S_w and maximizes the between-class scatter S_b . Hence we wish to find projection matrix that maximizes $\frac{|\phi^T S_b \phi|}{|\phi^T S_w \phi|}$. Let such a matrix be ϕ_{lda} .

In other words we need to find the maximum solution of the following equation

$$\begin{aligned}
|\Lambda| &= \frac{|\phi^T S_b \phi|}{|\phi^T S_w \phi|} \\
\Rightarrow \phi^T S_b \phi - \phi^T S_w \phi \Lambda &= 0 \\
\Rightarrow \phi^T (S_b \phi - S_w \phi \Lambda) &= 0 \\
\Rightarrow S_b \phi - S_w \phi \Lambda &= 0 \\
\Rightarrow S_b \phi &= S_w \phi \Lambda \\
\Rightarrow S_w^{-1} S_b \phi &= \phi \Lambda
\end{aligned}$$

Therefore, ϕ_{lda} will have eigenvalues of $S_w^{-1} S_b$ as its column vectors. It is a standard result in mathematics that eigenvalues of such matrices is at most $n - 1$. Hence the projection space of ϕ_{lda} is at most an $n - 1$ dimensional space.

Therefore, we have now successfully found a lower dimensional object, the projection space of ϕ_{lda} that is optimum for separating the classes. After projecting the sample data points onto this new dimensional object, one is now in a position to classify new data points by using different distance measures, like Euclidean Distance. Briefly, one looks at the distance of the new point from various classes and then allots it to the class which is the closest.

2.4.3 Classification accuracy

Classification accuracy is a way to calculate the percentage of correct predictions. The Figure 2.5 shows a confusion matrix which describes the performance of a supervised machine learning classification algorithm. Here, a is true positive, c is false positive, d is false negative and b is true negative. The classification accuracy says how often the classifier is correct. Sensitivity of the classifier model is given by $\frac{a}{a+d}$. The specificity of the model is given by $\frac{c}{c+b}$.

In our study, we are interested in classification accuracy since our aim is not to study the classifiers. Further, the classification accuracy value is correlated to the amount of information present in that particular set of features.

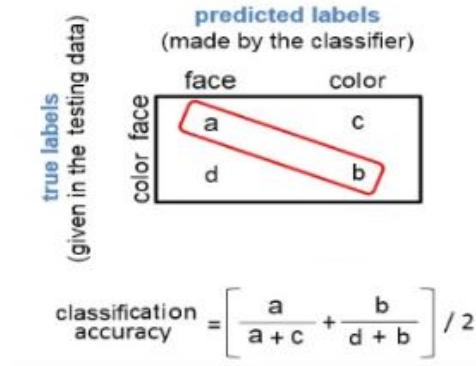


FIGURE 2.5: Classification accuracy.

2.5 Methods of Analysis

2.5.1 Region of Interest (ROI) based MVPA

The features for pattern analysis are the voxels that belong to anatomically / functionally defined regions which are of interest to the experimenter. MVPA is done on this region which results in a classification accuracy in terms of percentages which represent the task based information content in that region. For color and face perception tasks, the chosen ROIs are primary and secondary visual cortices, and ventral stream regions like V3v, V4v, LOC, FG for both scans (i.e., before and after practice sessions). Definition of ROIs included in this study relies on activation clusters obtained from SPM (Statistical Parametric Mapping) univariate analysis, done on the data.

We have used both LDA and SVM with leave-one-out cross validation method (n-1 runs are used for the training purpose and the remaining one run is used for testing) for analysis. To evaluate the classifier performance and its generalization across all the data, the cross-validation step was performed 8 times where each fold had a different run as the testing data and the classifier was trained on the remaining 7 runs. Then the accuracies of all cross-validation folds were averaged. This resulted in a single classification accuracy for each region of interest per subject.

2.5.1.1 Statistical significance test

For the group level results, statistical significance in the multivariate classification analyses was assessed using one sample t-test with the null hypothesis: "Classification done by the algorithms are just by chance, resulting in a mean classification accuracy of 50%". We reject the null hypothesis if the mean classification accuracy is significantly higher than the chance level at 5% significance level for a two-tailed test to prove the alternate hypothesis that the classification observed is not by chance but due to inherent differences in the patterns of the two classes [20].

To check if there is any significant difference in classification accuracies before practice and after practice, paired t-test was done for each ROI. The null hypothesis for the paired t-test: "There is no difference between the before and after practice images (the mean difference between paired observations is zero). We reject the null hypothesis if the mean difference is significantly higher than zero at 5% significance level.

2.5.2 Searchlight MVPA

A searchlight is like a repeated ROI analysis, where data in each searchlight can be described by a neighbourhood of features around a center feature. This approach can be applied equally to volume-based fMRI, surface-based fMRI. A searchlight map (or accuracy map) is created by applying an MVPA measure to data in each searchlight [21]. We have used a searchlight neighbourhood of 100 voxels at a time and LDA with leave-one-out cross-validation method for analysis.

The training and testing methods are similar to the methods followed in ROI analysis. However at the end of searchlight analysis, we get a new searchlight map for each subject per task where each voxel value is the classification accuracy of that particular searchlight having that voxel as the center feature. This new brain map can be used for further statistical significance tests at group level.

2.5.2.1 Statistical analysis

After obtaining the Accuracy Maps for all the 18 non-outlier subjects, we needed to get an idea of which voxels were statistically better than the others at classification. In order to do so we calculated the one-sample t-value of each voxel against the null hypothesis of it being a chance event (i.e. 50% or 0.5 value), where t-value is given by :

$$t = \frac{\bar{x} - 0.5}{\sigma/\sqrt{18}}$$

Here, \bar{x} references to mean classification accuracy of the particular voxel searchlight across 20 subjects and σ is its standard deviation.

Hence, we obtained t-value map of the brain i.e. a map of the voxels where each voxel is assigned its corresponding t-value. Then for each ROI, we compared it with the corresponding values in the t-value map to verify the ROI results.

2.6 Software and Tools

2.6.1 CoSMoMVPA

CoSMoMVPA is a multi-variate pattern analysis (MVPA) toolbox in Matlab [22]. MVP analysis techniques for any kind of neuroimaging data are supported by CoSMoMVPA. These toolboxes are used to address questions which are both data driven and hypothesis driven on neural organization and cognitive representations. The toolbox is very powerful and accommodates data analysis within and across space, time, neuroimaging modalities and organisms.

2.6.2 Statistical Parametric Mapping (SPM)

Statistical Parametric Mapping [23] is a process used to statistically test hypotheses on functional imaging data. A specialized software called SPM has been created for

this purpose. The inputs for SPM could be brain imaging data sequences from different groups of people, or time-series from the same subject. We have used SPM for pre-processing our data and for defining the regions of interest to create brain mask files.

2.6.3 MRICRON

MRICRON is a GUI based MRI image visualization toolbox which we have used for visualizing the output brain maps [\[24\]](#).

Chapter 3

Results

Both LDA and SVM classifiers performed equally well for this data set.

3.1 ROI Analyses

6 relevant functionally defined ROIs (also called "masks") have been chosen for the analysis. ROIs are similar but slightly different for both the scanning sessions since they are functionally defined. Different ROIs are of different sizes. Hence they span different number of slices. The 6 ROIs are:

1. Activated regions of left ventral stream (lven_act).
2. Activated regions of right ventral stream (rven_act).
3. Deactivated regions of left ventral stream (lven_deact).
4. Deactivated regions of right ventral stream (rven_deact).
5. Activated regions of primary and secondary visual regions in left hemisphere (lv1v2_act).
6. Activated regions of primary and secondary visual regions in right hemisphere (rv1v2_act).

Figure 3.1 shows all the relevant regions of interest.

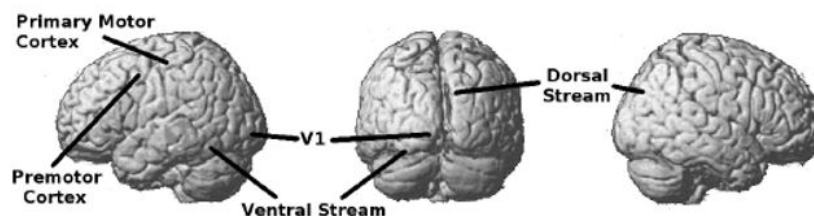


FIGURE 3.1: Regions of interest in the brain.

We have a combined mask for primary and secondary visual regions (V1)

The group level ROI analysis shows that all the ROIs significantly (significance level $p=0.05$) classify colour / rest (Figure 3.2 and 3.3) and face/rest (Figure 3.4 and 3.5) conditions for both before practice session (scan 1) and after practice session (scan 2) in both the classification methods (SVM and LDA).

The t-test results are given below:

Masks	SVM					
	Scan 1			Scan 2		
	Mean	Std. Dev.	P-values	Mean	Std. Dev.	P-values
lven_act	0.758	0.039	1.13E-15	0.688	0.088	5.83E-08
rven_act	0.704	0.061	8.11E-11	0.702	0.072	1.24E-09
lven_deact	0.581	0.044	4.07E-07	0.588	0.097	0.00125
rven_deact	0.635	0.058	2.18E-08	0.589	0.095	0.00098
lv1v2_act	0.693	0.1	2.82E-07	0.67	0.102	1.86E-06
rv1v2_act	0.727	0.079	8.76E-10	0.73	0.072	1.58E-10

FIGURE 3.2: Significance test for the SVM classification on colour perception.

Masks	LDA					
	Scan 1			Scan 2		
	Mean	Std. Dev.	P-values	Mean	Std. Dev.	P-values
lven_act	0.774	0.038	3.56E-16	0.696	0.083	1.53E-08
rven_act	0.704	0.061	8.11E-11	0.712	0.07	3.18E-10
lven_deact	0.602	0.047	5.73E-08	0.608	0.101	0.0002867
rven_deact	0.637	0.054	6.11E-09	0.61	0.112	0.0006617
lv1v2_act	0.696	0.101	2.34E-07	0.676	0.094	4.40E-07
rv1v2_act	0.737	0.086	1.65E-09	0.739	0.08	4.76E-10

FIGURE 3.3: Significance test for the LDA classification on colour perception.

To see the effect of practice on classification accuracies, we performed a paired t-test between before (scan1) and after (scan2) sessions. The only ROI which showed significant (significance level $p=0.05$) was the activated region of left ventral stream (lven_act), for colour perception. After practice the classification accuracy reduced significantly.

Masks	SVM					
	Scan 1			Scan 2		
	Mean	Std. Dev.	P-values	Mean	Std. Dev.	P-values
lven_act	0.746	0.072	5.65E-11	0.715	0.036	6.69E-15
rven_act	0.756	0.047	2.95E-14	0.709	0.063	9.43E-11
lven_deact	0.714	0.049	1.05E-12	0.669	0.036	3.74E-13
rven_deact	0.696	0.036	2.45E-14	0.658	0.042	1.00E-11
lv1v2_act	0.724	0.095	1.58E-08	0.716	0.081	2.51E-09
rv1v2_act	0.775	0.038	2.92E-16	0.779	0.062	7.28E-13

FIGURE 3.4: Significance test for the SVM classification on face perception.

Masks	SVM					
	Scan 1			Scan 2		
	Mean	Std. Dev.	P-values	Mean	Std. Dev.	P-values
lven_act	0.746	0.072	5.65E-11	0.715	0.036	6.69E-15
rven_act	0.756	0.047	2.95E-14	0.709	0.063	9.43E-11
lven_deact	0.714	0.049	1.05E-12	0.669	0.036	3.74E-13
rven_deact	0.696	0.036	2.45E-14	0.658	0.042	1.00E-11
lv1v2_act	0.724	0.095	1.58E-08	0.716	0.081	2.51E-09
rv1v2_act	0.775	0.038	2.92E-16	0.779	0.062	7.28E-13

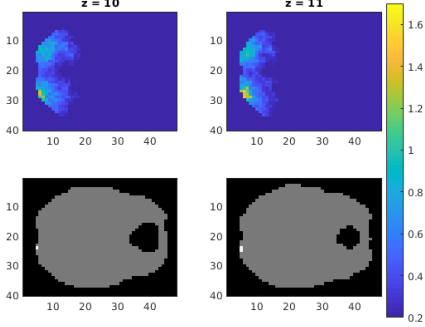
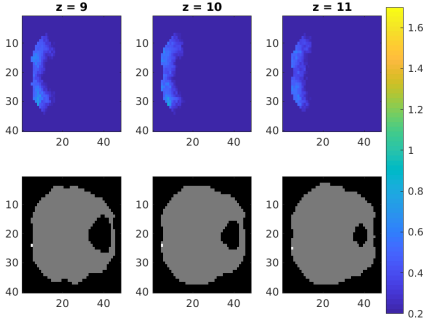
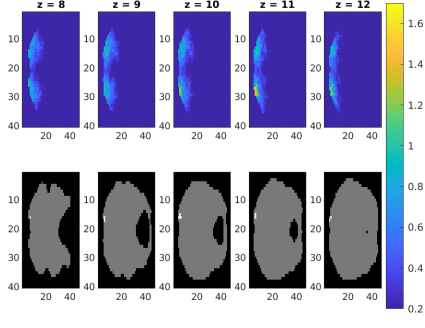
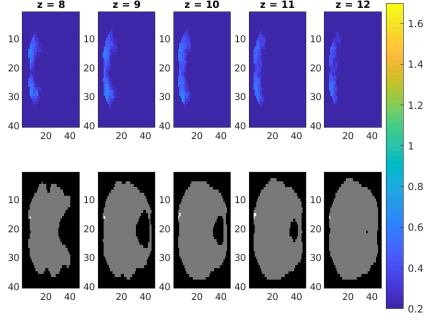
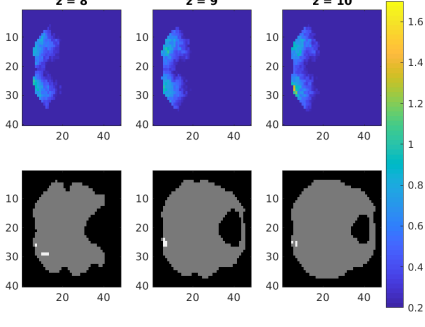
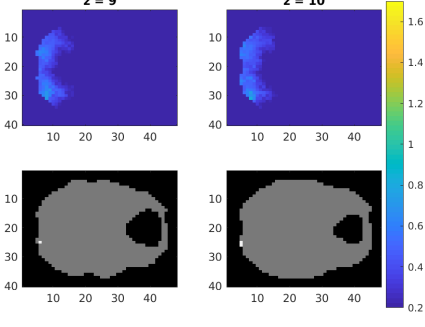
FIGURE 3.5: Significance test for the LDA classification on face perception.

3.2 Searchlight Analyses

The searchlight results re-affirm the role of ROIs in color and face perception. The t-value maps obtained from the searchlight analysis allows the identification of clusters across the whole brain which are statistically significant for task based information encoding.

The following images (Figure 3.6 and Figure 3.7) show the slices of t-value maps corresponding to the various ROIs. The ROIs themselves have been depicted in the second row as white regions in the brain mask (shown in grey color). The color coding for the t-values is show in the color bar present with each image. These images were created using MATLAB.

Color Perception

ROI	Before Practice	After Practice
lv1v2_act		
rv1v2_act		
lven_act		

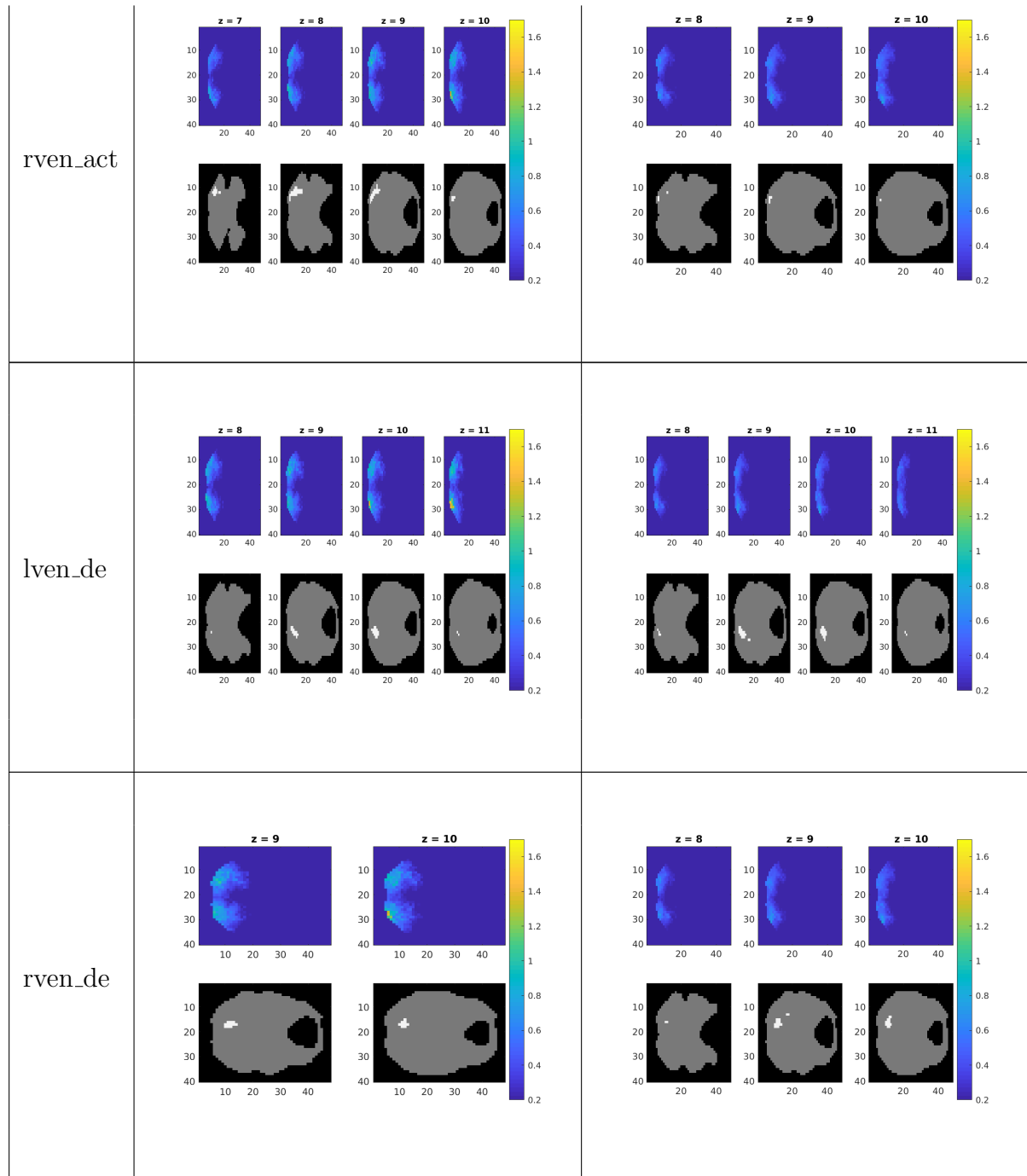
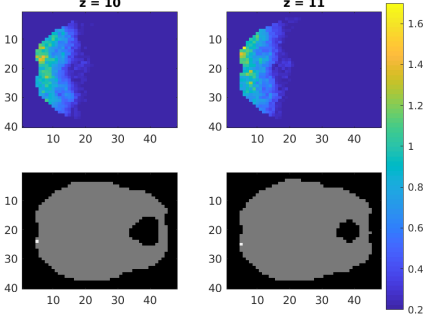
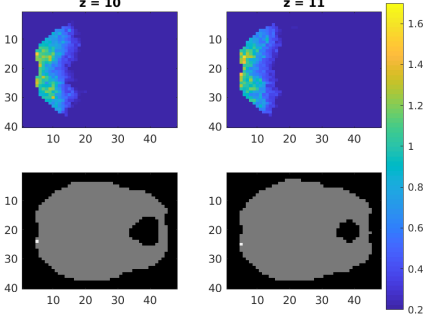
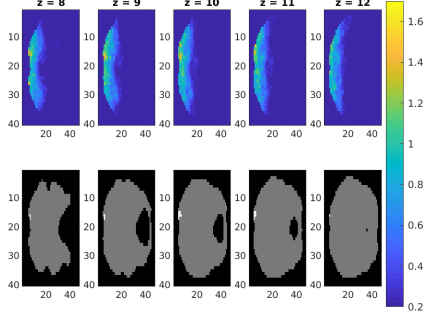
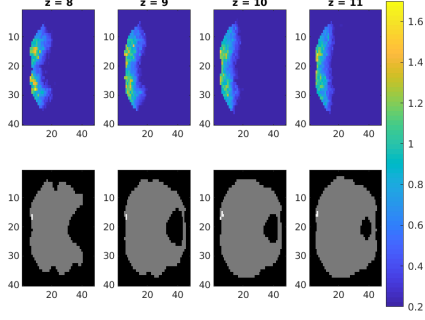
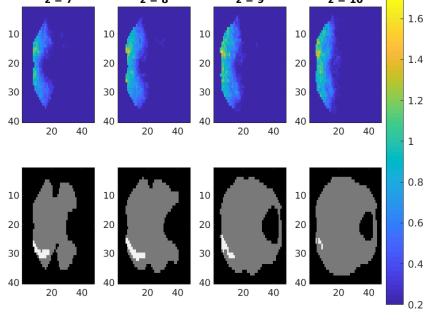
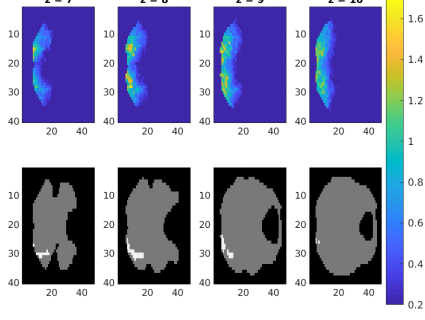


FIGURE 3.6: Color Perception

Face Perception

ROI	Before Practice	After Practice
lv1v2_act		
rv1v2_act		
lven_act		

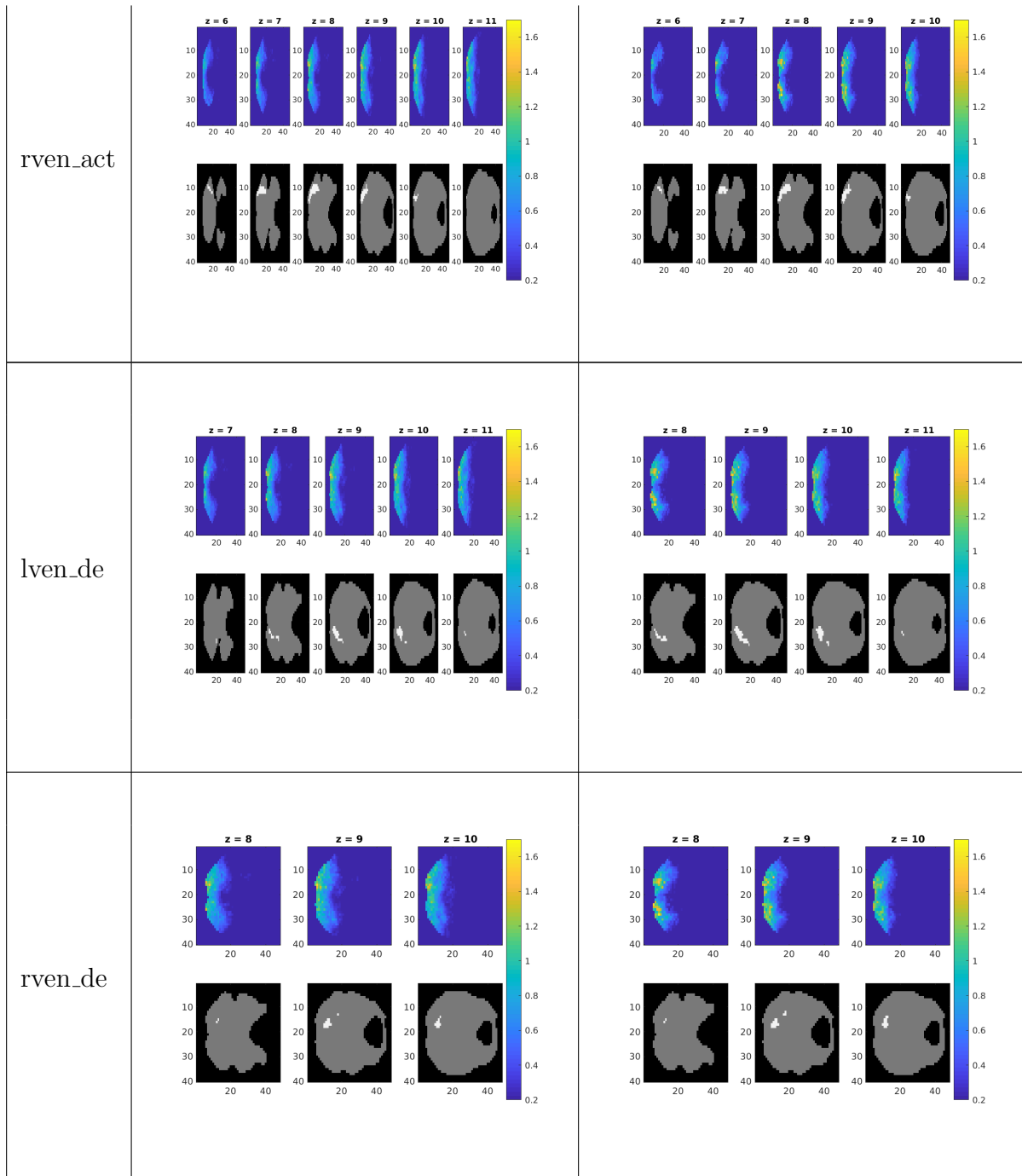


FIGURE 3.7: Face Perception

Chapter 4

Discussion and Future Direction

4.1 Possible Explanation for the Results

4.1.1 ROI Analysis

The six ROIs significantly classified color v/s rest and face v/s rest, which could possibly mean that those regions encoded task relevant information about colors and faces.

After practice, only the left ventral stream showed a reduced classification accuracy for the color perception task. A probable reason could be that region evolved to increase its efficiency of color recognition by restricting its stimulus dimensions.

4.2 Interpreting classification accuracy

Substantial advances have been made since the last decade, in fMRI decoding. However, it is very crucial to look at the challenges and short-comings in the analysis and interpretation of fMRI experiments. In our case, it is important to critically analyse classification experiments.

We know that fMRI is done using the BOLD signal which is a result of sluggish, nonlinear hemodynamic response. It is important to remember that BOLD signal is a proxy to neuronal signal and we are not capturing the neuronal activity directly. So, interpretations of our classification results are not directly addressing the information encoded in individual neurons [25].

4.2.1 Underestimation

Since, a voxel contains a pool of hundreds of neurons, it is possible to have a random mixture of neurons with different tuning properties. This may confound the macroscopic effect at the level of voxels. However, this does not mean that there is no information in the local neuronal populations. Also, a single neuron might encode a significant amount of information which is vanquished by the surrounding neurons which are contributing only noise. Thus the inherent nature of fMRI offers several instances to *underestimate* the information content in the brain. The hidden information can be unravelled only through direct invasive measurements of population signals combined with computational methods [17].

4.2.2 Overestimation

A voxel can also sample a huge blood vessel which supplies blood to a large population of neurons which are not anatomically related. This could result in a false positive result showing significant information in that region which is not computationally encoded by the neurons in reality.

The temporal resolution of fMRI is very low compared to the timescales of neuronal signal processing. The sluggish hemodynamic response which has a time lag of 1-2 seconds contributes to the temporal disintegration of neuronal signal and the image acquisition.

Thus there are several ways in which an observed fMRI classification accuracy might *overestimate* the information content at the neuronal level [17].

4.2.3 Other limitations

Following are few other reasons why the interpretations of classification accuracy can be challenging.

The design of the experiment, parameter setting, temporal aggregation, etc., can have huge impacts on the classification accuracies. The size of the data and the partitioning of the data-set into training and testing sets will decide how well the algorithm learns the optimal decision boundary.

Moreover, if we fit a very complex classification algorithm to the training set, we might face problems in generalization of the results to the testing set. This is called *over fitting* which is a major topic of concern in classification studies [17].

4.3 Conclusions

To put it in a nutshell, if the experiment and analysis is done with proper care and diligence, accounting for the errors in methodology, MVPA can reveal the mental representations encoded in the fMR signals. This is a breakthrough when compared to the classical univariate analysis.

MVPA is promising since it provides a generic framework for using fMRI data to validate theoretical / computational models in future. However, fMRI has limitations that are preventing us from extracting information from neuronal levels. So, the much needed move for a great leap is to corroborate the classification studies using neuronal recordings over different methods and different species [17].

4.4 What next?

MVPA is well known to be highly sensitive in detecting various characteristics of the stimulus. As we have seen, one of the methods of MVPA is the searchlight analysis. Traditional Searchlight analyses involved defining a sphere volume about

each voxel and associating with that voxel the classification accuracy of that whole sphere, thereby allowing us to project the experimental conditions of interest onto the brain, via the voxels. This is the analysis we are doing in our study. However, a major drawback of this analysis is that it does not take into account the actual structure of the brain i.e. the anatomy of the cortical surface [26].

A method that does account for this is called the **SURFACE-BASED SEARCH-LIGHT** analysis, in which we reconstruct the cortical surface using the anatomical information of each subject. This surface is now used to select the region around a voxel whose accuracy will be associated with this voxel. This approach differs from volume based searchlight analysis in two aspects :

1. Since we use voxels on the reconstructed cortical surfaces we are restricting our analysis to a more realistic subspace namely the grey matter of our brain which contains all the cell bodies.
2. We use geodesic metric as a distance measure instead of the standard 3D-euclidean metric. This will avoid selection of false neighbours (voxels) across the sulcus.

Clearly, it is interesting and crucial to understand the effect of these two factors on the actual classification accuracy. At the same time it is necessary to decode if the information obtained is sensitive to the spatial structure. Since, surface-based analysis provides a better representation of the information-containing part of our cerebral cortex, it might give a better understanding of the color and face perception regions of the brain.

Bibliography

- [1] *Cognition*. Psychology Today. URL: <https://www.psychologytoday.com/basics/cognition> (visited on 04/15/2019).
- [2] *Functional neuroimaging*. ScienceDaily. URL: https://www.sciencedaily.com/terms/functional_neuroimaging.htm (visited on 04/15/2019).
- [3] S. Sandrone et al. “Angelo Mosso(1846-1910)”. In: *Journal of Neurology* 259 (Aug. 23, 2012), pp. 2513–2514. DOI: <https://doi.org/10.1007/s00415-012-6632-1>. URL: <https://link.springer.com/article/10.1007%5C%2Fs00415-012-6632-1>.
- [4] *Foundation Text — Neuroimaging: Visualizing Brain Structure and Function*. URL: <http://ccnmtl.columbia.edu/projects/neuroethics/module1/foundationtext/> (visited on 04/15/2019).
- [5] OpenStax. *Version 8.25 from the Textbook*. May 18, 2016. URL: <https://commons.wikimedia.org/w/index.php?curid=30147912> (visited on 04/15/2019).
- [6] *Read Principles of fMRI — Leanpub*. URL: <https://leanpub.com/principlesoffmri/read> (visited on 04/16/2019).
- [7] *Multi-voxel Pattern Analysis (MVPA) and Mind Reading By: James Melrose. - ppt download*. URL: <https://slideplayer.com/slide/5212144/> (visited on 04/15/2019).
- [8] Selket. *English: Image showing dorsal stream (green) and ventral stream (purple) in the human brain visual system*. URL: https://commons.wikimedia.org/wiki/File:Ventral-dorsal_streams.svg (visited on 04/16/2019).

- [9] D. Ray et al. “Visual Dual Stream is Input-based, Neuroplastic and Predictive”. In: *Unpublished*. Visual Dual Stream is Input-based, Neuroplastic and Predictive ().
- [10] Goodale. MA and Milner. AD. “Separate visual pathways for perception and action”. In: *Trends in Neuroscience* 15 (1992), pp. 20–25. ISSN: 1. DOI: [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8).
- [11] E. Darcy Burgund and Chad J. Marsolek. “Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems”. In: *Psychonomic Bulletin & Review* 7.3 (Sept. 1, 2000), pp. 480–489. ISSN: 1531-5320. DOI: 10.3758/BF03214360. URL: <https://doi.org/10.3758/BF03214360> (visited on 04/15/2019).
- [12] James T. Enns. *The Thinking Eye, the Seeing Brain* — W. W. Norton & Company. URL: <https://books.wwnorton.com/books/webad.aspx?id=10584> (visited on 04/15/2019).
- [13] *Perception Lecture Notes: Recognition*. URL: <http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/recognition/recognition.html> (visited on 04/15/2019).
- [14] Joel Z Leibo, Jim Mutch, and Tomaso Poggio. “Why The Brain Separates Face Recognition From Object Recognition”. In: (), p. 9.
- [15] A.C.. Evans et al. “3D statistical neuroanatomical models from 305 MRI volumes”. In: *IEEE* 4939523 (Aug. 6, 2002). DOI: 10.1109/NSSMIC.1993.373602.
- [16] Tyler Davis et al. “What Do Differences Between Multi-voxel and Univariate Analysis Mean? How Subject-, Voxel-, and Trial-level Variance Impact fMRI Analysis”. In: *NeuroImage* 97 (Aug. 15, 2014), pp. 271–283. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2014.04.037. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4115449/> (visited on 04/15/2019).
- [17] John-Dylan Haynes. “A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives”. In: *Neuron* 87.2 (July 15, 2015), pp. 257–270. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2015.05.025. URL: [http:](http://)

- [//www.sciencedirect.com/science/article/pii/S0896627315004328](http://www.sciencedirect.com/science/article/pii/S0896627315004328) (visited on 04/15/2019).
- [18] James V. Haxby. “Multivariate pattern analysis of fMRI: The early beginnings”. In: *NeuroImage*. 20 YEARS OF fMRI 62.2 (Aug. 15, 2012), pp. 852–855. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2012.03.016. URL: <http://www.sciencedirect.com/science/article/pii/S1053811912002893> (visited on 04/15/2019).
- [19] *Linear Discriminant Analysis*. Assignment Point. May 22, 2016. URL: <http://www.assignmentpoint.com/business/business-statistics/linear-discriminant-analysis.html> (visited on 04/16/2019).
- [20] Tae.K. Kim. “T test as a parametric statistic”. In: *Korean Journal of Anesthesiology* (Nov. 25, 2015). DOI: 10.4097/kjae.2015.68.6.540.
- [21] Joset A. Etzel, Jeffrey M. Zacks, and Todd S. Braver. “Searchlight analysis: Promise, pitfalls, and potential”. In: *NeuroImage* 78 (Sept. 1, 2013), pp. 261–269. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2013.03.041. URL: <http://www.sciencedirect.com/science/article/pii/S1053811913002917> (visited on 04/15/2019).
- [22] Nikolaas N. Oosterhof, Andrew C. Connolly, and James V. Haxby. “CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave”. In: *Frontiers in Neuroinformatics* 10 (2016). ISSN: 1662-5196. DOI: 10.3389/fninf.2016.00027. URL: <https://www.frontiersin.org/articles/10.3389/fninf.2016.00027/full> (visited on 04/15/2019).
- [23] K.J. Friston et al. “Statistical Parametric Mapping: The Analysis of Functional Brain Images. Elsevier, London.” In: (2006).
- [24] Chris Rorden, Hans-Otto Karnath, and Leonardo Bonilha. “Improving lesion-symptom mapping”. In: *Journal of Cognitive Neuroscience* 19 (2007), pp. 1081–1088. DOI: <https://doi.org/10.1162/jocn.2007.19.7.1081>.
- [25] N. K. Logothetis et al. “Neurophysiological investigation of the basis of the fMRI signal”. In: *Nature* 412.6843 (July 12, 2001), pp. 150–157. ISSN: 0028-0836. DOI: 10.1038/35084005.

- [26] Nikolaas N. Oosterhof et al. “A comparison of volume-based and surface-based multi-voxel pattern analysis”. In: *NeuroImage*. Multivariate Decoding and Brain Reading 56.2 (May 15, 2011), pp. 593–600. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2010.04.270. URL: <http://www.sciencedirect.com/science/article/pii/S1053811910006981> (visited on 04/15/2019).