UNIVERISTY OF TARTU

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Institute of Computer Science
Information Technology Curriculum

Martin Loginov

Beyond decoding: representational similarity analysis on fMRI data

Master's Thesis (30 EAP)

Supervisor: Raul Vicente Zafra, PhD

Tartu 2015

Beyond decoding: representational similarity analysis on fMRI data

Abstract:

Representational similarity analysis is a novel data analysis technique in neuroscience first proposed by Kriegeskorte et al. in [KMB08]. It aims to connect different branches of neuroscience by providing a framework for comparing activity patterns in the brain that represent some cognitive processes. These activity patterns can come from various sources, like different subjects, species or modalities like electroencephalography (EEG) or functional magnetic resonance imaging (fMRI). The central concept of RSA lies in measuring the similarity between these activity patterns. One of the open questions regarding RSA is what distance measures are best suited for measuring the similarity between activation patterns in neuronal or fMRI data.

In this thesis RSA is implemented on a well known fMRI dataset in neuroscience, that was produced by a studying the categorical representations of objects in the ventral temporal cortex of human subjects [HGF+01]. We carry out RSA on this dataset using different notions of distance and give an overview of how the end results of the analysis are affected by each distance notion. In total 9 different distance measures were evaluated for calculating the similarity between activation patterns in fMRI data. The results provided in this thesis can be used by researchers leveraging RSA to select distance measures for their studies that are most relevant to their particular research questions at hand. In addition to the comparison of distance notions, we also present a novel use case for RSA as a tool to visualize the global effects different transformations can have on the input dataset.

Keywords: representational similarity analysis, fMRI, ventral temporal cortex, distance measures

Dekodeerimise tagamaad: fMRI andmete esituste sarnasuse analüüs

Lühikokkuvõte:
Andmeesituste sarnasuse analüüs on uudne andmeanalüüsimeetod neuroteaduste kontekstis, mis pakuti välja Kriegeskorte et al. poolt artiklis [KMB08]. Selle eesmärgiks on ühendada erinevaid neuroteaduste harusid, luues ühtne raamistik erinevaid kognitiivseid protsesse esindavate aktiivsusmustrite omavaheliseks võrdlemiseks ajus. Raamistik võimaldab omavahel võrrelda tegevusmustreid, mis pärinevad erinevatest allikatest, nagu erinevatelt katsealustelt, liikidelt või on mõõdetud kasutades erinevaid tehnoloogiaid nagu elektroentsefalograafia (EEG) või funktsionaalse magnetresonantstomograafi (fMRI) abi. Andmeesituste sarnasuse analüüsi keskne idee seisneb nende mustrite omavahelise sarnasuse võrdlemises. Üks lahtisi küsimusi selles valdkonnas seisneb sobivate mõõdikute leidmises funktsionaalse magnetresonantstomograafi poolt mõõdetud aktiivsusmustrite omavahelise sarnasue hindamiseks.

Käesolev magistritöö viib läbi andmeesituste sarnanuse analüüsi ühe neuroteadustes teada tuntud fMRI andmestiku peal [HGF$^+$01]. Analüüsi käigus uuritakse erinevate kaugusmõõdikute mõju andmeesituste sarnasuse analüüsi lõpptulemustele ja antakse sellest põhjalik ülevaade. Kokku on uurimise all 9 erinevat kaugusmõõdikut. Käesoleva töö tulemusena valmib põhjalik ülevaade erinevate kaugusmõõdikute mõjust andmeesituste sarnasuse analüüsile, mida on võimalik kasutada edasistes neuroteaduse alastes uurngutes hindamaks erinevate kaugusmõõdikute sobivust mõne konkreetse uurimustöö kontekstis. Lisaks kaugusmõõdikute võrdlemisele pakub käesolev töö välja ka ühe uudse kasutusjuhu andmeesituste sarnasuse analüüsile. Nimelt on võimalik kõnealust meetodit kasutada ka erinevate lähteandmetele rakendatavate teisenduste mõju visuaalseks hindamiseks.

Võtmesõnad: admeesituste sarnasuse analüüs, funktsionaalne magnetresonantstomograafia, kaugusmõõdikud

Contents

1 Introduction

The human brain is one of the most complex biological structures in existence. The inner workings of this incredible organ, that is responsible for every single thought or complex cognitive process in humans and animals alike, have for long been the subject of scientific study in many different fields. One on the key questions in neuroscience is determining how the brain represents different types of information. Studies have shown that different areas in the brain are associated with different types of information. There are specific regions associated with the processing of low level visual information (Vu et al. [VRN+11]), while other areas process information on a more abstract level, dealing with the categorization of objects (O'Toole et al. [OJAH05]; Haxby et al. [HGF+01]).

While it is known that information processing differs between brain regions, little is known about the exact nature of information representation in these different regions or how a set or neurons maintain information. To alleviate some of the shortcomings of existing analysis techniques in neuroscience, a novel method was proposed in a paper by Kriegeskorte et al. called *representational similarity analysis* [KMB08]. RSA provides a robust framework for studying how different stimuli or cognitive processes are represented in the brain. It has since been used to study how the categorical representation of objects differs between humans and monkeys for example [KMR+08].

While RSA is showing great promise in connecting different branches of neuroscience, there are also numerous questions still open. The central concept of RSA lies in measuring the similarity between activity patterns in the brain. These activity patterns can be measured using different technologies, for example by directly measuring the electrical activity from the brain using electroencephalography (EEG) or by indirectly measuring blood flow using functional magnetic resonance imaging (fMRI). In both cases the activity patterns themselves are composed of information coming from different channels (in the case of EEG) or different voxels (for fMRI) - they are multivariate. One of the core questions raised in a recent review article by Kriegeskorte et al. is: "What multivariate distance is best suited for measuring representational dissimilarities in neuronal or fMRI data?" [KK13]

This thesis aims to study the effects of different notions of distance in measuring the similarity between activity patterns that make up neural representations in the context of fMRI data. Our goal is to assess how the different distance measures influence the results of representational similarity analysis for activation patterns in fMRI data. In total we evaluate nine different notions of distance by implementing representational similarity analysis with each of them on a well known fMRI dataset in neuroscience [HGF+01]. The results presented here can be used to enhance the utility of RSA as a analysis technique in neuroscience. By providing a thorough overview of the effects resulting from different distance measures we provide researchers the ability to select for their analysis the distance measures that are most relevant to their particular research questions.

The bulk of this thesis is organized very coarsely into four core chapters:

- Chapter 2 provides some background to readers completely unfamiliar with neuroimaging techniques. This is done to provide a better understanding of the dataset used in this thesis.

- Chapter 3 gives a thorough overview of all the different steps involved in the analysis for this thesis. It starts with the description of the dataset and preprocessing steps, familiarizes the reader with the concepts of classification analysis and rep-

resentational similarity analysis and ends with the description of all the different notions of distance we evaluated in this thesis.

- Chapter 4 contains all the results obtained from different analysis stages. The reader is presented with the results from classification analysis, followed an example representational similarity analysis. The bulk of this chapter consists of the results we obtained testing different notions of distance. In the end we also present a novel use case for representational similarity analysis we discovered while doing research for the thesis.

- Chapter 5 tries to present some possible interpretations for the results obtained in this thesis and the reasoning behind them. Some ideas for further research are also proposed.

2 Background information

This section tries to provide some background information to a reader completely unfamiliar with brain imaging techniques. The main goal here is not to provide a thorough technical overview, but rather convey some intuition about what can be accomplished with different brain imaging techniques and how they can be used for measuring brain activity.

2.1 What is MRI

MRI stands for magnetic resonance imaging and is a technique for creating detailed images of organs and other types of soft tissue inside an organism. An MRI scanner works by applying a very strong magnetic field to the atoms making up the different tissues. "The magnetic field inside the scanner affects the magnetic nuclei of atoms. Normally atomic nuclei are randomly oriented but under the influence of a magnetic field the nuclei become aligned with the direction of the field. The stronger the field the greater the degree of alignment. When pointing in the same direction, the tiny magnetic signals from individual nuclei add up coherently resulting in a signal that is large enough to measure." [Dev07]

The magnitude of this summary signal varies between different types of tissue - this is the key concept behind MRI technology that facilitates the creation of detailed anatomical images. When used for brain scans, this variability in signal strength makes it possible to differentiate between between gray matter, white matter and cerebral spinal fluid in structural images of the brain. You can see what an MRI scan looks like on Figure 1.
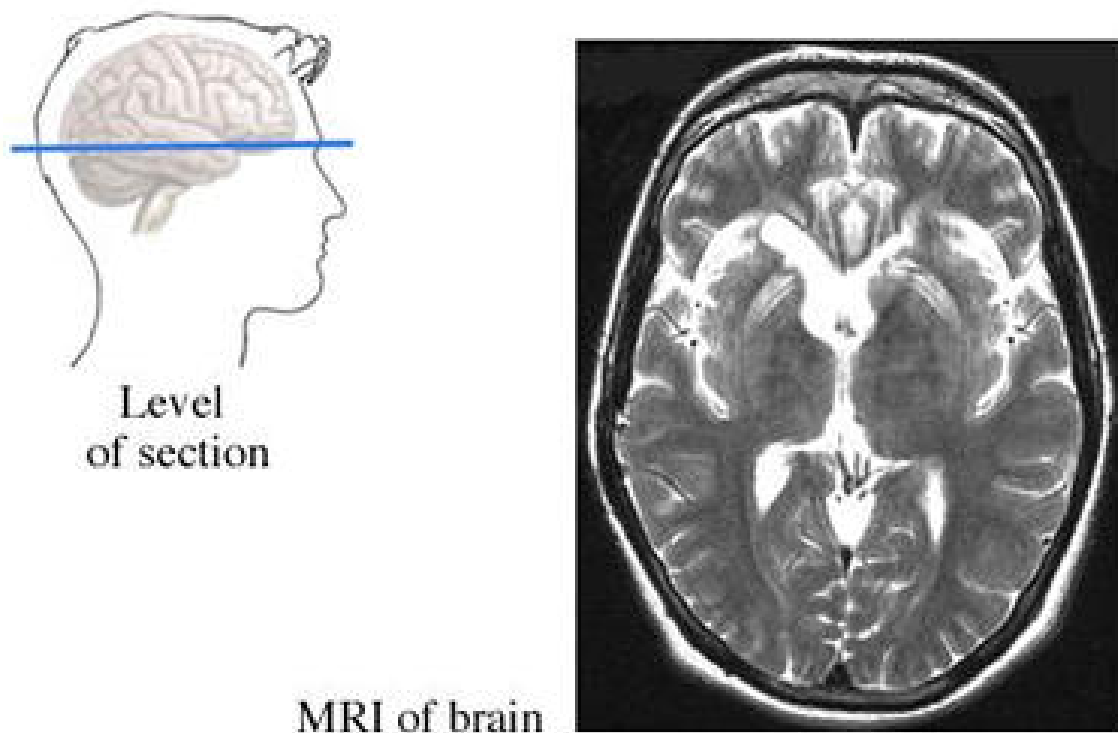

Figure 1: Single slice from a brain MRI scan [Slo13]

The image obtained as a result of a scan is actually 3-dimensional. This point is

illustrated on Figure 1 as it shows only a single slice cut through the brain and there are many such slices on different levels.

The entire image is made up of tiny units called voxels [1]. Depending on the scanner and scan type the size of a voxel can vary from less than a cubic millimetre to a couple of cubic millimetres, therefore encompassing brain tissue consisting of millions of cells. The size of a voxel effectively determines the resolution of the resulting image. The higher the resolution, the more time it takes for the scanner to create a full brain image. This limitation becomes important when taking functional images (see Section 2.2) and is usually the reason why functional images are created with a lower resolution than anatomical images.

2.2   fMRI

While MRI facilitates the creation of detailed images of brain anatomy, fMRI (functional magnetic resonance imaging) provides the ability to indirectly measure brain activity. Functional imaging is an enhancement to the regular MRI technique to account for blood flow inside the brain.

The underpinnings of modern fMRI were first discovered by Seiji Ogawa. While conducting research on rat brains, he was trying out various different settings on his MRI scanner and discovered a contrast mechanism reflecting the blood oxygen level of different brain tissue [Log03]. Today, Ogawa's discovery is known as blood oxygen level-dependent signal (BOLD). It is also known that BOLD contrast depends not only on blood oxygenation but also on cerebral blood flow and volume.

These changes in blood flow, volume and oxygenation, that are linked to some neural activity are known as the hemodynamic signal. Cognitive processes in the brain consist of millions of neurons firing. Unlike electroencephalography (EEG) which measures electrical activity from the neurons more or less directly, the hemodynamic BOLD signal is a indirect measure of neural activity. This thesis does not concern itself with the exact neuroscientific semantics of the link between the BOLD signal and brain activity. For our intents and purposes we just consider the BOLD signal to be influenced by neural activity in a way that a change in the signal in interpreted as a change in neural activity.

The data from fMRI experiments is a set of 3D volumes consisting of voxels. Each voxel has an intensity value that represents the BOLD signal in that particular area of the brain. During a scan session hundreds of such 3D volumes are recorded, usually while the subject is performing some task. These volumes together form a timeseries and represent the change in the hemodynamic signal in time and are the subject for analysis in this thesis.

---

[1]A voxel or volumetric pixel can be thought of as the 3-dimensional equivalent to a 2-dimensional pixel.

3   Methods

This chapter provides a detailed overview of all the data analysis techniques used in this thesis. For starters we describe how the data was obtained, the experiment that produced it, followed by a detailed description of preprocessing steps to make it viable for statistical analysis. Before describing the core concepts behind classification analysis and representational similarity analysis, we also give an overview how and why we localized the core analysis steps to a particular region in the brain. The chapter ends with the descriptions of all the different notions of distances we evaluated for this thesis.

Readers not particularly interested in all the technical details can skip this chapter and go straight to chapter 4 that contains all the results from the different analysis steps described in this chapter. Relevant details from this chapter are also referenced from there.

3.1   Obtaining the data

The data used in this thesis was acquired from the OpenfMRI project [PBM+13] database. OpenfMRI aims to provide researchers a common platform and infrastructure to make their neuroimaging data freely available to any interested party.

The project tries to solve some of the big challenges with sharing fMRI datasets existing today, like the large size of datasets and lack of standardization in data organization and preprocessing steps. By providing a common format to organize different types of data (imaging data, experimental task descriptions, behavioural data) and tools for preprocessing the data, they are greatly simplifying the process of obtaining datasets for neuroscientific studies.

3.2   The dataset: neural responses to grayscale images

The specific dataset we used was produced for a study by Haxby et al. [HGF+01], where they studied the representations of faces and objects in the ventral temporal cortex of human subjects. This section will give a thorough overview of their experiment and the data produced.

During the experiment neural responses were collected from six subjects (five female and one male) using a General Electric 3T fMRI scanner while they were performing a one-back repetition detection task. In essence subjects were presented with different stimuli consisting of gray-scale images that depicted objects belonging to different categories (see Figure 2) and for each stimulus image the subjects had to indicate whether the image was the same or different as the previous image.

By design the experiment was a block design experiment. This is one of the two common types of experiments used to gather fMRI data (the other being event-related design) [AB06]. In a block design experiment two or more conditions are being presented to the subject in alternating blocks and in each block only one condition is active. During each block a certain number of fMRI scans are taken and these scan volumes will represent the neural responses for the condition at hand. The blocks are usually separated by a period of rest to allow the hemodynamic signal to return to a baseline. This makes it easier and more robust to distinguish the signals representing different conditions.

In this specific experiment the condition for a block was the category of a stimulus image. Stimuli were gray-scale images of faces, houses, cats, bottles, scissors, shoes,
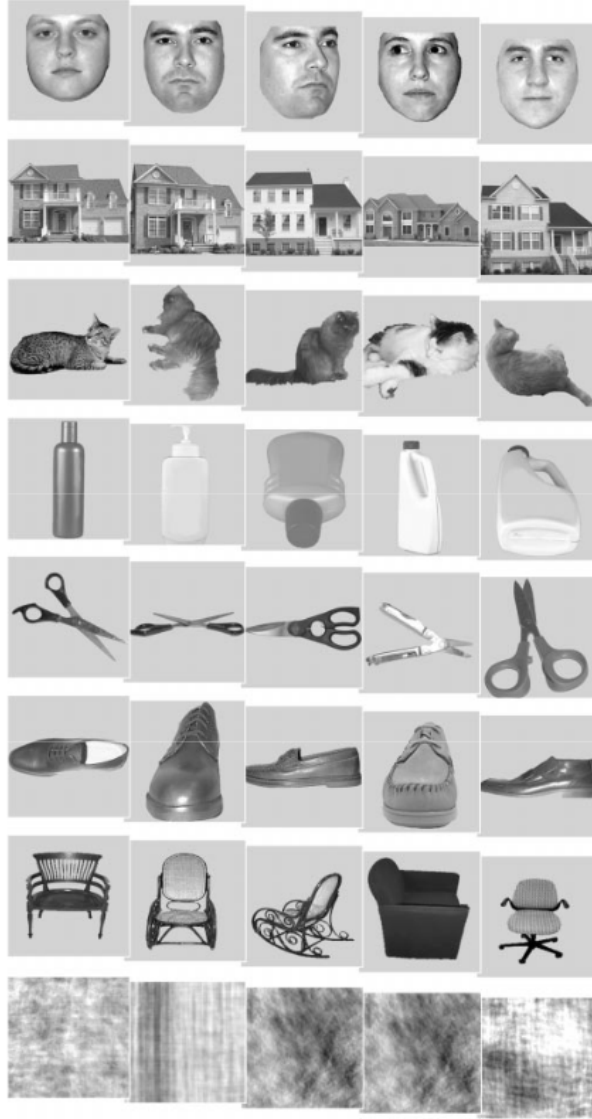
Figure 2: Examples of stimuli. Subjects performed a one-back repetition detection task in which repetitions of meaningful pictures were different views of the same face or object. [HGF+01]

chairs and nonsense patterns - eight different categories in total. Twelve time series were obtained for each subject. Each time series began and ended with 12s of rest and contained eight stimulus blocks of 24-s duration, one for each category, separated by 12-s intervals of rest. An illustration of a single time series (a singe run of the experiment) is shown on Figure 3.

Stimuli were presented for 500 ms with and interstimulus interval of 1500 ms. Repetitions of meaningful stimuli were pictures of the same face or object photographed from different angles. Stimuli for each meaningful category were four images each of 12 different exemplars.

For each timeseries 121 3-dimensional fMRI volumes were obtained with dimensions of 40 x 64 x 64 voxels. With the scan parameters used[2] each voxel represents the neural

---

[2]repetition time (TR)=2500 ms, 0 3.5-mm-thick sagittal images, field of view (FOV) = 24 cm, echo time (TE) = 30 ms, flip angle = 90
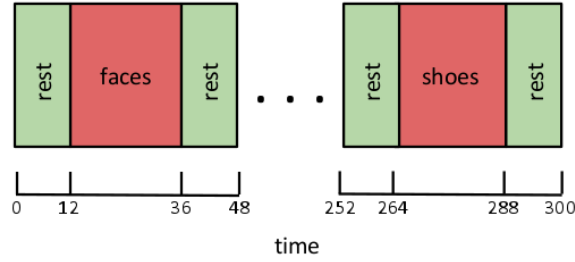
Figure 3: Depiction of the different stimulus blocks in a single time series of the experiment. Time is shown on the horizontal axis from the start of the run.

activity in an area of brain tissue the size of 3.5 x 3.75 x 3.75 mm. The count of volumes for each series (121) comes from the fact that under these scan parameters (specifically the repetition time (TR) = 2500 ms) it takes 2.5 seconds to carry out a full brain scan. The last scan is initiated after the final 12s of rest in each timeseries (after timepoint 300s).

Each of the scan volumes is a representation of either one of the active experiment conditions (category of a stimulus) or the resting state and is labeled accordingly. The voxel values in these scan volumes together with the labels serve as input to analysis in this thesis.

It might seem surprising to the reader that nothing is done with the results of the one-back repetition detection task - in fact this kind of behavioural data is not even included in the OpenfMRI database for this dataset. The main goal of the one-back task is actually to trick the subjects into concentrating on the individual images which would allow them also to perceive the category of the image. Without this kind of task at hand the attention of the subjects would likely dissipate at least once during the five minute test run and this kind of background activity from wandering thoughts would add additional noise to the scan results.

3.3   Software used for processing the data

Right from the beginning of this project we decided to use IPython Notebooks as our main analysis platform [PG07]. Having the ability to run arbitrary Python code on remote servers conveniently through a web browser proved invaluable to us when the memory and computation requirements of the analysis exceeded the capacity of a single laptop. Python's large ecosystem of 3rd party libraries coupled with IPython Notebook's ability to visualize plots in-line with the actual code and notes make it a very productive and accessible environment for exploratory data analysis.

The other main reason besides accessibility for choosing this analysis platform was the existence of the PyMVPA toolbox (MultiVariate Pattern Analysis in Python). The authors describe it as: "a Python package intended to ease statistical learning analyses of large datasets. It offers an extensible framework with a high-level interface to a broad range of algorithms for classification, regression, feature selection, data import and export." [HHS+09]

PyMVPA abstracts away much of the complexities associated with analyzing neuroimaging data. It provides functions for directly loading in datasets that are structured according to the OpenfMRI specification [PBM+13] and provides numerous means for

preprocessing them. Scaffolding for running cross-validated classification and representational similarity analysis is also provided, which is leveraged extensively by the analysis steps in this thesis.

A similar toolbox to PyMVPA exists that is targeted specifically to carrying out representation similarity analyses in Matlab, as described in a paper by Nili et al. [NWW+14a]. Although the paper describes it to be somewhat similar with respect to RSA functionality, it lacks the additional features and flexibility of PyMVPA. In addition, Matlab is a proprietary platform which would have required us to deal with licensing issues.

Besides PyMVPA another notable piece of software in our analysis pipeline is the Scikit-Learn machine learning library for Python [PVG+11]. It is a library that provides simple and efficient tools for data mining and data analysis. Besides the fact that PyMVPA uses it internally for classification analyses for example, we also use it directly to carry out dimensionality reduction leveraging the results of representational similarity analysis.

For visualizing the raw fMRI data the MRIcron toolkit was used [RKB12]. This toolkit was used to visualize region of interest maps and the results of our experiments with searchlight analysis (see Section 3.5).

3.4   Preparing the fMRI data for analysis

The output of most fMRI experiments is a number of 3D scan volumes (images from now on) for each run of the experiment. Each image just contains the raw BOLD intensities for each voxel and represents the neural activity for the entire scan area (usually the whole brain) at a particular timepoint. To make any meaningful statistical analysis on these images viable, a number of preprocessing steps need to occur to transform the raw data.

This section gives an overview of the preprocessing steps carried out on the dataset to prepare it for analysis. We divide the steps into two: some are "standard" preprocessing techniques, that are performed on most fMRI datasets before analysis and are in fact already applied for all datasets in the OpenfMRI database, the other steps we applied ourselves with respect to our particular analysis. As the technical details for the former are beyond the scope of this thesis, we only give a list of the more important steps for the sake of completeness and describe each step briefly in terms of its input and output, for more thorough explanations refer to [Str06]. Some of the more important preprocessing steps in the order they are implemented:

1. **Conversion of file formats:** as data can come from different scanners that use different formats, everything is converted into a common format - NIfTI (Neuroimaging Informatics Technology Initiative file format [CAB+04]). This is done to facilitate inter-operation of different functional MRI data analysis software.

2. **Motion correction:** subjects inevitably make small movements during experiments. This causes the signal-to-noise ratio of the resulting scan images to drop. Motion correction is used to calculate a set of rigid body transformations for each image to counteract this.

3. **Brain extraction:** the scanner itself has no means to distinguish between the brain and other tissue types. Brain extraction is later used as a preprocessing step to extract from the images an image mask representing only the voxels that contain actual brain tissue.

The above list is not meant to be a comprehensive enumeration of all the steps that went into preprocessing the raw data, it is merely there to give the reader a glimpse of different types of methods that need to be used before any higher level analysis can occur. Preprocessing fMRI data is a complex field - entire university courses are taught on the subject. Next we describe a set of steps taken during the analysis of this thesis and add some reasoning for each step.

Since the spacial structure of images is not really needed in any parts of the analysis (after applying region of interest masks as described in Section 3.5) all the images are converted into vectors of voxel intensities by simply concatenating all the voxel values in a volume:

$$\textbf{image} = [v_1, v_2...v_n]$$

Here v_1 through v_n are the individual voxel intensities. Representing the images like this makes working with the data much simpler because it is much clearer to reason in terms of vectors for most data analysis algorithms, than it would be for 3D volumes.

Even after motion correction, the data can still contain some global effects, like low frequency signal intensity drift in the intensity values of voxels throughout a timeseries. This phenomenon is described in a technical paper by Tanabe et al. where they say that two known sources of drift are noise from the MR scanner and aliasing of physiological pulsations [TMT+02]. The paper centers around the issue that there is not yet a clear consensus on which algorithm should be used for removing these sorts of drifts and presents the comparison of a few methods.

In this thesis we used a technique called *detrending*, provided by the PyMVPA toolkit [HHS+09]. The main idea behind detrending is modeling the drifts in intensities as polynomial trends. It basically transforms the data by fitting polynomials to the timeseries of each voxel, leaving only information that is not explained by the polynomials. In the simple case of linear detrending, a straight line is fit through the timeseries of a voxel using linear regression and the voxel initensity values are replaced with the residuals from the regression.

After detrending we would like to increase the signal to noise ratio for the data even more by trying to focus on the parts of the signal that are related to the stimuli. For this we transform the timeseries of each voxel so that it no longer contains the raw BOLD signal intensities, but a deviation from a baseline signal. The idea behind this is that we only want to focus on activity elicited by a stimulus, not on activity that is there all the time. In our case, the resting state will serve as the baseline. This technique is called z-scoring, although the classical version differs from the variation presented here:

1. The timeseries of each voxel is considered separately from all the others. A timeseries for a voxel is composed of all the values for this single voxel for a specific run of the experiment.

2. Calculate the mean intensity value of the voxels for resting state only.

3. Calculate the standard deviation of intensity values of the voxels for resting state only.

4. For all the voxels in a timeseries, calculate the deviation from the mean calculated in step 2 in units of standard deviation, step 3.

After z-scoring, we remove all images for resting state from the dataset. This is done because resting state images are not interesting (nor particularly useful after z-scoring has

13

been carried out), since all subsequent analysis stages deal only with the representations of the different stimuli.

The final output from the preprocessing stage for a single subject comprises of 864 image vectors each containing the the voxel intensities of about 500 voxels, as defined by the region of interest. This number of images takes into account images from all the 12 time series with resting state images removed.

3.5   Region of interest - ventral temporal cortex

In our dataset each fMRI volume in a timeseries contains over 160 000 voxels that make up the neural representation of the entire brain at a single point in time. Processing this amount of data is not actually a problem with the computational resources readily available today, although some types of analysis using these full brain images can still take a long time, for example classification analysis. Still we confine all subsequent analysis implemented in this thesis to only a particular region in the brain, for reasons that will be explained in the following paragraphs.

At all times, even during seemingly restful periods, there are always some background cognitive processes happening in the brain that cause the firing of millions of neurons. These firings cause changes in the hemodynamic BOLD signal that is measured by the fMRI scanner. Some of these cognitive processes can be localized to specific areas in the brain [PPFR88] while others seem to occur globally. In addition it is known that the representation of the same information can differ greatly between different brain regions[NPDH06]. As we are interested in only the representations of categories of visual objects, activity from all these other processes become noise in the dataset with respect to our goals. Therefore we would like to localize our analysis to a region in the brain that we know contains the neural representations of the particular conditions or phenomena that are of interest to us - this is known as a region of interest in neuroscience. This sort of localization is important because the data coming from the fMRI scanner is inherently noisy, distinguishing our signal from all the other processes in the entire brain would be very difficult if not impossible.

A ROI can be defined either anatomically or functionally. An anatomically defined region is an area in the brain that is known from previous studies or research to contain the neural representation of interest. In practice for fMRI datasets, regions of interest are defined by masks. A mask represents the notion that from all the voxels in the brain, we select only the ones belonging to our ROI and mask the others. Masks are carefully drawn for each subject (as each brain is unique to a certain extent) by neuroscientists using landmarks in the brain that are common to all animals in a species. In this thesis the region of interest was also chosen anatomically to be the ventral temporal cortex in the brain, shown on Figure 4.

The ventral temporal cortex in humans has been extensively studied and is known to represent visual stimuli in an abstract manner that allows these stimuli to be grouped into different categories. Some studies that have shown this categorical organization of representations include [OJAH05] and [HMH04] among others. For the dataset used in this thesis the ventral temporal cortex consists of 577 voxels.

During research for this thesis we also tried to define our region of interest functionally. This begins by first hypothesising that there is a specific area in the brain where the information of interest (categories of objects) is represented in a way that is relevant to our analysis. In our case we would like that the representations of different cate-
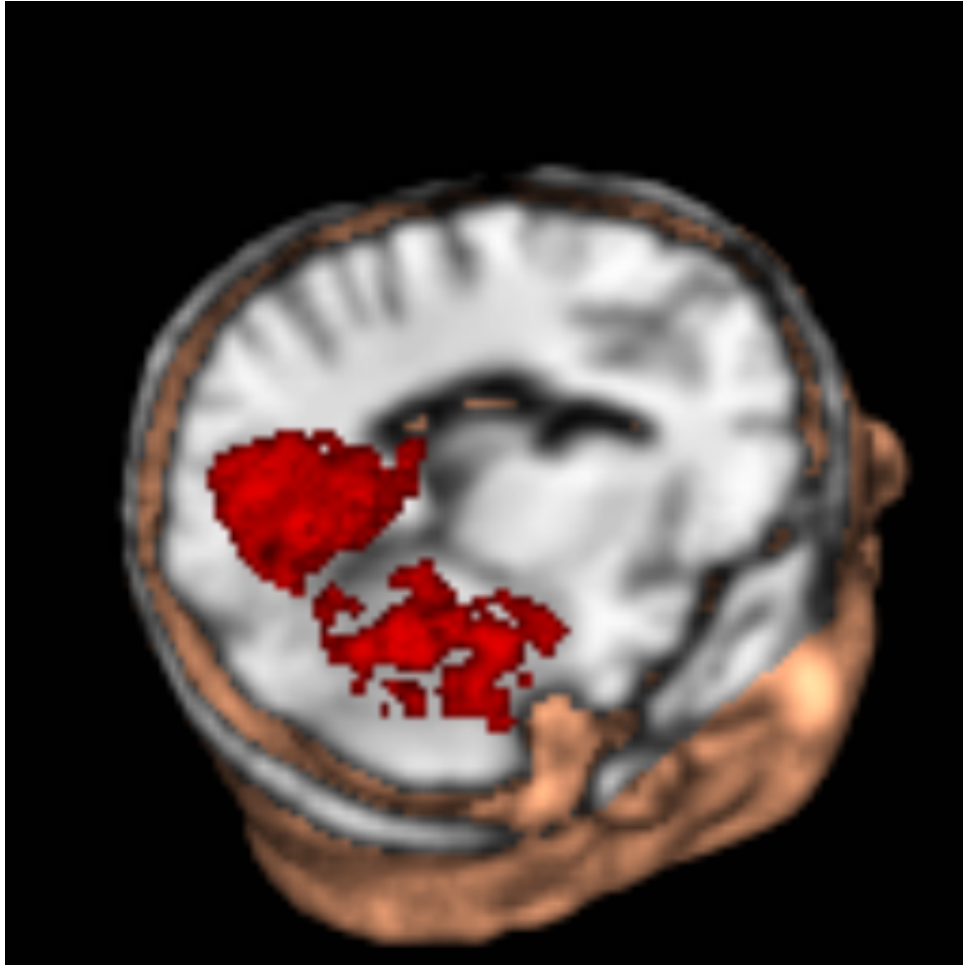
Figure 4: A cutout from the anatomical fMRI volume of a subject, voxels from the ventral temporal cortex are colored red. The figure was rendered using the MRIcron tool suite for processing fMRI data [RKB12].

gories are clearly separable from each other. To find this region we used a technique called searchlight analysis, as proposed by Nikolaus Kriegeskorte and Peter Bandettini in [KB07]. Searchlight analysis takes into account the spacial structure of fMRI volumes and is composed of the following steps:

1. **Define a radius as a number of voxels.** This radius will determine the size of the area (in voxels) that is considered for each iteration in the analysis.

2. **Iterate over all the voxels in the dataset.** In our case we have $40 * 64 * 64 = 163840$ for a single fMRI volume. On each iteration we only consider voxels that are in the immediate vicinity of the current voxel at hand, as defined by the radius in step 1. This essentially allows us to move through the entire brain and localize all our processing to a small area in each iteration. In fact the term *searchlight analysis* originates from the notion that in essence we are looking for something in the brain by casting a figurative searchlight (our region defined by the radius) around the entire brain and examining things we can observe in the light.

3. **Calculate a metric for each region to assess the relevancy of this particular area for the research question at hand.** In our case we ran classification analysis as described in Section 3.6 to determine how well the representations of different categories are separated in a specific region. A high classification accuracy in identifying fMRI volumes that contain the representation of a particular stimulus category would indicate the suitability for further analysis of this region.

As a result of searchlight analysis, the region we found (with the highest overall classification accuracy in identifying different categories) was to a large extent overlapping with the ventral temporal cortex defined anatomically. Since the average classification accuracy in this functional region was only marginally higher than in the ventral temporal cortex, we opted to use the anatomically defined region for all subsequent analysis instead.

This decision in favor of using the anatomical region was influenced by two factors. First there are numerous pitfalls and limitations in interpreting the results from regions defined by searchlight analysis in a neuroscientifically meaningful way. The discussion of these limitations is beyond the scope of this thesis, but is covered by Etzel et al. in [EZB13]. Second, since data from the ventral temporal cortex in the same dataset we used is already extensively studied, as mentioned above, we wanted our results to be directly comparable with the previous studies. Defining a different region would have made direct comparisons more difficult while at the same time giving us no additional benefits.

3.6 Classification analysis

As a result from the preprocessing stage we obtained 864 vectors with fMRI image data. These vectors each contain the neural representation of a particular stimulus category and are labeled accordingly. Before we can go on to study the representations directly, we have to make sure that this dataset contains enough information to distinguish between representations of different categories in the experiment. In other words we would like to know if images for different categories would still be distinguishable from each other if they were not labeled.

For this purpose we train a classifier. From a very high level perspective a classifier is a function that is trained to identify samples from a dataset based on some features.

In our particular case the samples are the images and the features are individual voxel values. A set of samples together with their label are used as input to train the classifier (Figure 5).
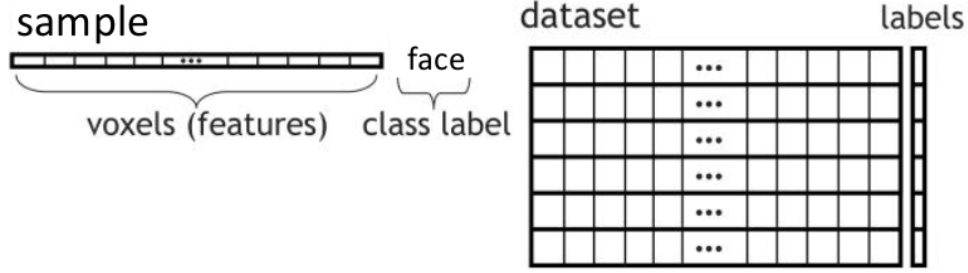

Figure 5: Depiction of a dataset consisting of samples used for the training procedure of a classfier. Figure adapted from [PMB09].

There are many different types of classifiers and they all work differently, but their general concept of operation is similar. They contain some internal parameters which are tuned according to the training dataset. These parameters are supposed to capture the underlying structure of the data based on the features. The idea is that if a classifier manages to capture the underlying structure in the data, then it can use it to predict labels for samples it has never seen before. Given a sample $\mathbf{x} = [x_1...x_n]$ the classifier f would predict it's label y:

$$y = f(\mathbf{x})$$

To assess the quality of a classifier, it needs to be tested on a different set of samples from the ones used to train it. "The typical assumption for classifier learning algorithms is that the training (and testing) examples are independently drawn from an 'example distribution'; when judging a classifier on a test set we are obtaining an estimate of its performance on any test set from the same distribution [PMB09]." The process of training and evaluation of a classifier is depicted on Figure 6.
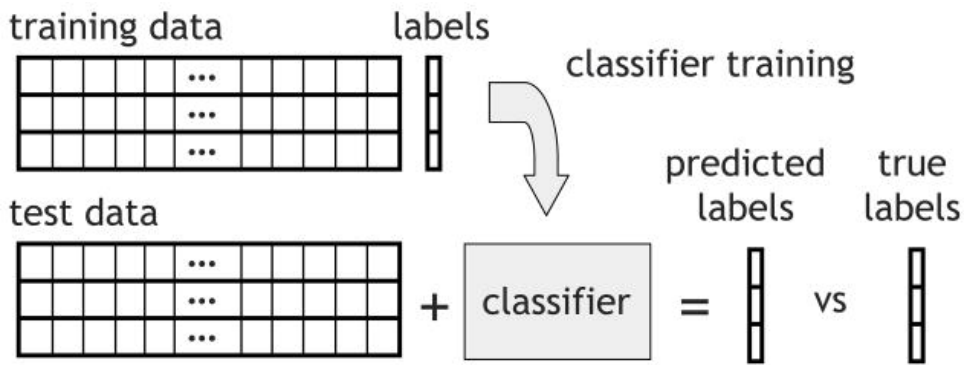

Figure 6: Classifier training and evaluation [PMB09].

To actually measure the quality of a classifier function we calculate a metric called *classification accuracy*:

$$\frac{\sum_{i=1}^{n} I(f(\mathbf{x_i}), y_i)}{n} \qquad (1)$$

Here n is the number of samples in the test set, $\mathbf{x_i}$ and y_i are the ith sample vector and label from the test set respectively and I is a function returning 1 if $f(\mathbf{x_i}) = y_i$ and 0 otherwise. In other words classification accuracy is the ratio of correctly classified samples from the test set to all the samples in the test set.

Partitioning data samples into training and test sets for validation can be done in a multitude of ways. Ideally we would like to have as many samples as possible for training the classifier to make it more accurate, while still having a test set that is a good representative sample of the entire dataset. In order to achieve this we opted for a training/test procedure called *cross-validation.*

The idea behind cross-validation (also sometimes called *n-fold cross-validation*) is to partition the entire dataset into n folds. It is imperative that each fold contains the same proportion of samples from all the different classes, to help the classifier capture the structure of the different samples as equally as possible. After partitioning the following procedure is implemented, which is also illustrated on Figure 7:

1. Leave one of the folds out and train the classifier using samples from all the other $n-1$ folds. Use the data from the left out fold for testing.

2. Repeat step 1 for each fold in turn.

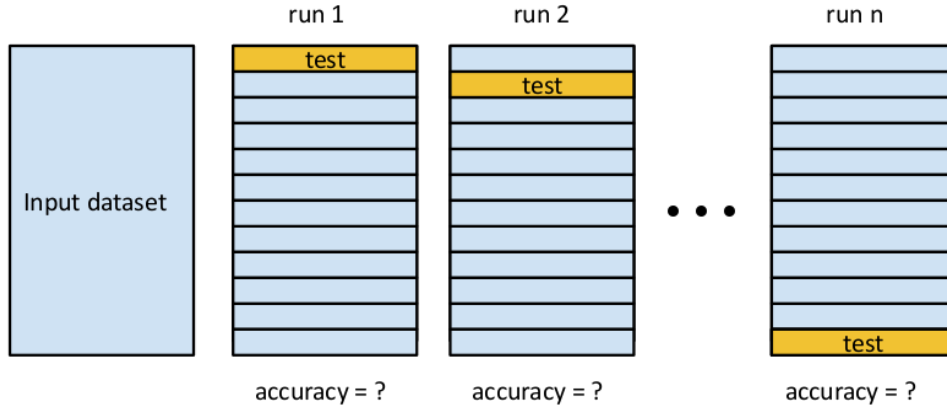3. Calculate the accuracy for each of the folds used in testing.

Figure 7: Illustration of the n-fold cross-validation procedure, depicting the partitioning of an input dataset into separate folds and one fold on each run being used for testing.

After this the average accuracy across all folds can be calculated by:

$$\sum_{i=1}^{n_{folds}} \frac{C_i}{n_{folds}} \tag{2}$$

Here C_i is the accuracy of each individual run in the cross-validation procedure as calculated by equation (1). This average accuracy is a good indication of the overall expected classifier performance on all samples taken from the same distribution as the input dataset.

In our particular case partitioning our 864 sample images into folds turns out to be trivial. As we gather from the description of the experiment (Section 3.2) the data is already partitioned into different runs with each of the runs containing every stimulus block exactly once. Therefore when taking each run to be a fold in the cross-validation

procedure, we have to make no additional effort in making sure that each fold contains the correct proportion of examples from each class.

All classification analyses in this thesis were carried out using a C-SVM classifier with a linear kernel. The choice of classifier was based on no particular reason other that being readily available in the PyMVPA toolkit. For a more thorough discussion regarding linear SVM classifiers refer to [HDO$^+$98].

All 864 images for a subject were used as input for the cross-validation procedure, no block-wise or category-wise averaging was carried out on the sample images (as is common for block-design fMRI experiments). The reasoning behind this was that averaging would reduce the number of samples in the training dataset and reasonable classifiction performance cannot be expected when the training dataset contains less samples than there are features per sample. Even after extracting only voxels from a specific region of interest (see Section 3.5) the sample image vectors would still contain hundreds of voxels.

We did not spend any time tuning classifier parameters as classifiers in the context of this thesis are used only as a sort of litmus test for a particular set of fMRI images (considering data from only a specific region of interest). Good classifier accuracy at identifying the stimulus category that each image represents is a precondition for representational similarity analysis as it validates our assumption about different neural representations being distinguishable.

3.7 Representational similarity analysis

Representational similarity analysis is a novel data analysis framework in the context of neuroscience, first proposed by Kriegeskorte et al. in [KMB08]. In this section a thorough overview of RSA will be given, first describing some background by drawing comparison to other data analysis techniques in neuroscience. Then we elaborate on the different steps involved in RSA and finally describe the scope within which RSA is used in this thesis.

Classical analysis techniques in functional neuroimaging can very broadly be divided into two: univariate and multivariate analysis. In the context of fMRI univariate analysis deals with individual voxels, determining voxels that react maximally to a certain experimental condition for example. This is used to localize cognitive processes in the brain and also to determine whether some cognitive process correlates with a predefined model. Training classifiers on fMRI images is an example of multivariate analysis. Here cognitive states in the brain are modeled as a set of voxel activations usually in reaction to an experimental condition. This type of analysis is also known as decoding and can be used determine whether different experimental conditions are differentiable in a specific region of interest in the brain.

Both univariate and multivariate methods process some kind of measured activity patterns in the brain that represent cognitive states. These activity patterns can be either voxel intensities in the case of fMRI or voltage spikes from neuronal cell recordings in the case of electroencephalography (EEG). If we devised an experiment where for each experimental condition both fMRI and EEG data were recorded, then direct comparison of activity patterns between these two different modalities would likely be so difficult as to make it infeasible. Indeed it would require us to devise a correspondence mapping from a timeseries of voxel intensities to a timeseries of voltage spikes. The same problem exists even when comparing activity patterns from the same modality: fMRI recordings from different subjects are also not directly comparable, as the structure of each individual

brain differs enough that there is not a one to one correspondence between the voxels in two different brains.

RSA alleviates this problem with direct correspondence mapping by abstracting data away from different activity patterns into a representational space. The central notion in RSA is a representational dissimilarity matrix (RDM). This matrix encodes the similarity structure between different activity patterns which in turn represent different experimental conditions. By comparing RDMs instead of activity patterns directly, we are able to compare the representations of cognitive states not only between different subjects or species, but also between different modalities and even between experimental measures and computational models. This powerful concept is illustrated on Figure 8.
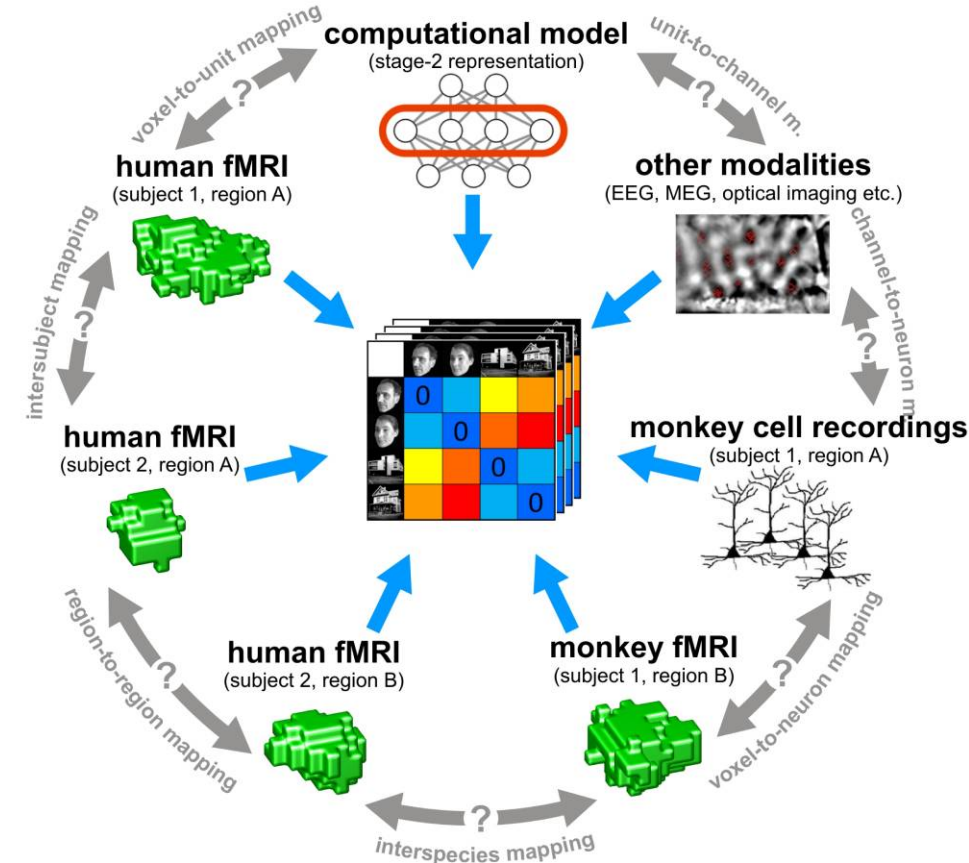

Figure 8: Depiction of how representational dissimilarity matrices facilitate the comparison of congnitive states between different subjects, species, modalities and regions in the brain [KMB08].

RSA as proposed in [KMB08] consists of five steps. In the following all the steps are presented together with descriptions and examples on how they are implemented in the context of the this thesis.

1. **Estimating the activity patterns.** The analysis starts by estimation of activity patterns for each experimental condition. In our case the activity patterns are voxel intensities from the ventral temporal cortex (see Section 3.5) as 864 image vectors.

   Since the experiment that produced the patterns was of block-design, there are some implications regarding how we can input our data for RSA. In a block-design experiment there exists a one to one correspondence between *all* the data from

a stimulus block and an experimental condition, but not for individual samples. What this means is that individual image vectors from a block do not contain all the information for representing the neural state for a given experimental condition (the category of a stimulus image in our case)[3].

We counteract this issue by averaging all the individual image vectors from one stimulus block together to act as a representation of this entire block. Recall that the experiment contained 12 runs for each subject and each run contained 8 blocks, therefore after averaging, instead of 864, we obtain 96 image vectors that each contain the neural representation of a single expermental condition.

2. **Measuring activity-pattern dissimilarity.** In this step we actually calculate the dissimilarities between the activity patterns representing different conditions. Between each pair of activity patterns a dissimilarity measure is calculated and together these values form a representational dissimilarity matrix (RDM). The RDM is a square matrix with the row and column length equaling the number of different experimental conditions. The matrix is symmetrical around a diagonal of zeroes (the dissimilarity between each condition and itself is 0).

In this thesis we use nine different methods to assess the similarity of activity patterns. All these different notions of distance are described in Section 3.10.

3. **Predicting representational similarity with a range of models.** Suppose we convert measured activity patterns from different regions in the brain (or different brains) into representational space by calculating RDMs. Although we are now able compare these representations to each other directly, we could still only assess whether the brains, that could even belong to different species, represented the same set of stimuli similarly or differently.

While this is already an achievement, the real utility of RSA lies in the fact that we are now able to relate representations from models to actual representations in the brain - this could give new insight into the inner workings of different areas in the brain.

As an example consider a computational model consisting of artificial neurons constructed specifically to mimic the proposed information processing occuring in some brain region. We can now feed the same set of stimuli to both the model and actual human subjects. Using RSA to compare the representations in both the real brain and inside the model, we can make reasonable assumptions about the information processing structure in the brain, if there is a high correlation between actual representations and model representations.

Models used for RSA do not have to be as complex as the hypothetical one described above. Models can also be constructed from behavioral data, for example reaction times to certain stimuli - basically anything that can be converted to RDM form, can act as a model in the context of RSA. In this thesis we use an even simpler type of model - a conceptual model, differentiating between animate and inanimate objects (see Section 4.2).

---

[3]This problem did not arise during classification analysis in Section 3.6, when we used the individual image vectors for training the classifier function. This comes from the fact that during training the classifier function accumulates information from the individual samples into internal parameters and because it eventually sees all the samples from a stimulus block we can reason that it therefore has the ability to "learn" the entire neural representation corresponding to an experimental condition.

4. **Comparing dissimilarity matrices**

   After calculating RDMs that encode the representation of different experimental conditions in either different regions of interest or models, they can be visually or quantitatively compared. For such comparisons we could measure the similarity of the RDMs themselves by calculating a dissimilarity matrix of dissimilarity matrices. We use this method in Section 4.3 for comparing RDMs obtained from the activation patterns with different notions of distance.

   To assess the similarity between two RDMs, we employ a measure called Kendall's tau (τ): it represents the proportion of pairs of values that are consistently ordered in both RDMs under comparison (see Section 3.10.10).

5. **Visualizing the similarity structure of representational dissimilarity matrices by MDS**

   To visualize the similarity structure contained in the RDMs, we can leverage multi-dimensional scaling. MDS is a general purpose dimensionality reduction algorithm for transforming datapoints inhabiting a high dimensional space to a much lower dimensional space (usually 2D or 3D) while at the same time trying to preserve the proportional distances between points.

   We use MDS to visualize the similarity between the representations of activity patterns estimated in step 1 and also to visualize the similarity between the RDMs of different distance notions in Section 4.3.

As we can witness from the description above, representational similarity analysis provides a powerful framework that allows us to represent the description of cognitive states in a more abstract space, providing much more flexibility compared to classical analysis methods and enabling new opportunities for relating datasets originating from different species, modalities and models.

3.8 Ordering representational dissimilarity matrices using hierarchical clustering

Gaining any meaningful insights about the data by the visualization of representational dissimilarity matrices alone is rather unlikely. This is especially true if the rows and columns of the RDM are randomly ordered. Different orderings of the rows and columns can be quite revealing however, but coming up with any kind of meaningful ordering requires some sort of a priori knowledge about the underlying similarity structure present in the dataset. In our case we could order rows and columns in the RDMs by the stimulus categories or by the experiment run number. In the first case it would group together all representations belonging to the same category and in the second case all the representations from the same experiment run. This would make sense, because we expect the representations of the same stimulus category to be more similar to each other than to representations of other categories. The reasoning is the same behind grouping by experiment run numbers.

The issue with this approach is that it would only facilitate the testing of existing hypotheses. In order to visualize the *real* similarity structure between representations, we reordered the rows and columns of RDMs using hierarchical clustering. The goal is to group together representations that are naturally similar to each other, without making assumptions about the underlying similarity structure of the activity patterns.

In general hierarchical clustering is used to identify clusters in datasets based on some distance notion between the points. In our case the datapoints are fMRI activity patterns representing the category of a stimulus image and we already have all the pairwise distances between our datapoints encapsulated in the representational dissimilarity matrix. The algorithm starts by finding the two closest points (with the smallest dissimilarity value in the RDM) and grouping them together. This group is now considered a single point in the dataset. On the next iteration all the pairwise distances between points are once again considered to find the next closest points. Depending on how the distances between groups of points are calculated the algorithm is said to perform either single, average or complete linkage clustering. In total $n-1$ such iterations are performed, where n is the number of points in the dataset.

While a thorough description of hierarchical clustering is not in the scope of this thesis, we refer the reader to chapter 7 of the book [Gre07]. The entire chapter centers around hierarchical clustering analysis and provides a worked example through the algorithm along with a discussion about different variations.

The output of hierarchical clustering is a dendrogram where the leaf nodes are our activity patterns and they are grouped together based on their similarity. The branches of the dendrogram leave a hierarchical trail of when in the process two nodes were connected. We used this ordering of the leaf nodes to reorder the rows and columns of the dissimilarity matrix. Figure 9 shows the RDM for the visualization of the similarity between different stimulus categories as calculated by the Mahalanobis distance (Section 3.10.9). The dendrogram is also visualized on the sides of the RDM to illustrate the process.

Just for comparison, the same RDM is shown on Figure 10, but there the rows and columns are randomly ordered. Clearly we can get more intuition about the underlying similarity structure between the stimulus categories from the visualization in Figure 9. We can immediately observe that small objects like shoes, bottles and scissors seem to have more similar representations than objects from other categories. This observation is not obvious by looking at Figure 10.

3.9   Dimensionality reduction using multidimensional scaling

Although we showed in Section 3.8 that representational dissimilarity matrices can be transformed to visualize the information they contain in a more meaningful way, they are in general not the best tool for visualizing the similarity structure between representations of cognitive states. A more useful technique for creating this kind of visualization is called multidimensional scaling. MDS is a general dimensionality reduction technique first proposed by Joseph Kruskal in a 1964 paper [Kru64]. Since then multiple variations of this classical version have been developed [BG05]. In this thesis we used the MDS implementation provided by the Scikit-Learn machine learning toolkit exclusively [PVG+11].

In general MDS is used to reduce the dimensionality of datapoints residing in a high dimensional space and project these points into a space of much lower dimensions. The real utility of MDS lies in the fact that this projection is carried out in such a way that distances between the points in the high dimensional space are preserved as much as possible in the lower dimensional space. The dimensionality of the space to project points onto is usually chosen to be either two or three dimensions to facilitate visualization.

We use MDS extensively to construct 2D scatterplots in order to visualize the similarity between representations of different stimulus categories in the brain. The representations are activation patterns consisting of voxel intensities and their dimensionality
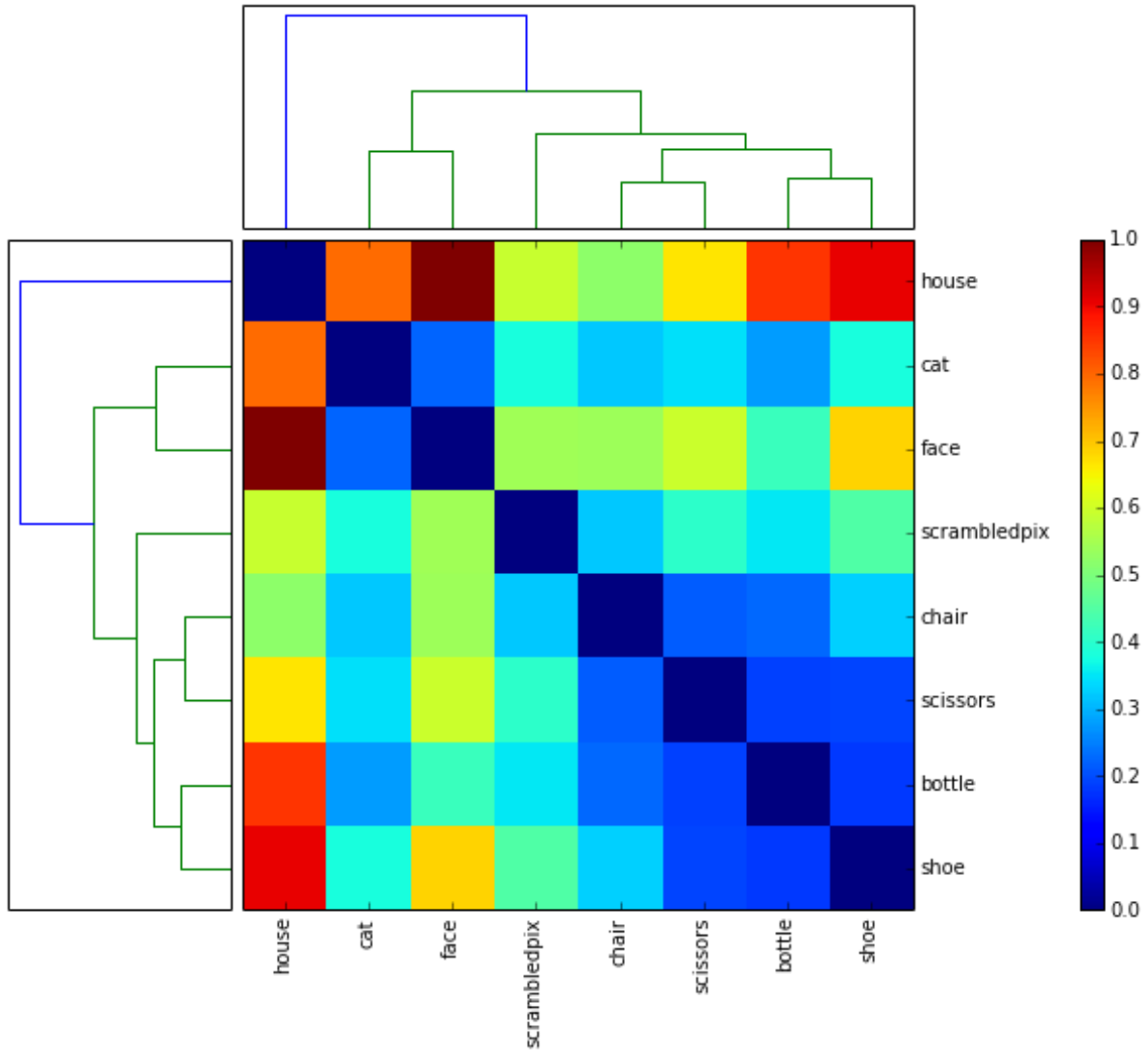
Figure 9: Representational dissimilarity matrix for the stimulus categories as calculated using the Mahalanobis distance. The rows and columns are ordered based on the ordering of the leaf nodes in the dendrogram resulting from hierarchical clustering.
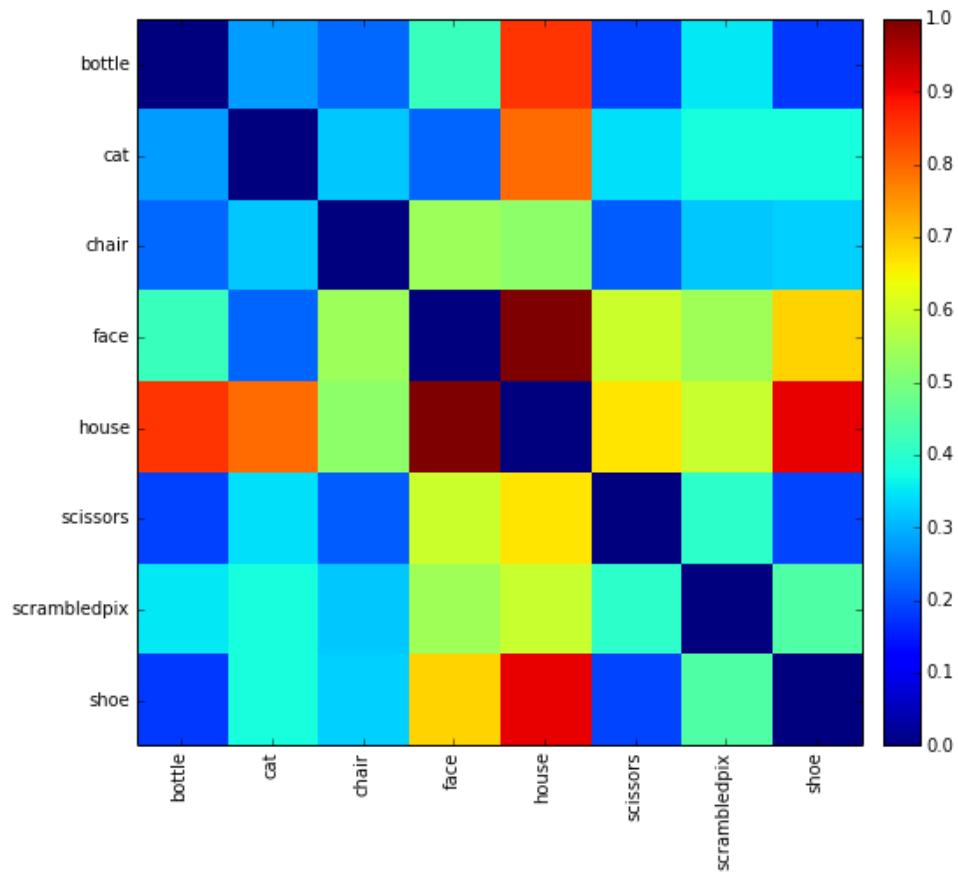
Figure 10: Representational dissimilarity matrix for the stimulus categories as calculated using the Mahalanobis distance. Rows and columns are randomly ordered.

is defined by the region of interest. Since the region of interest used in this thesis (the ventral temporal cortex 3.5) is the size of about 500 voxels in our scan data, we essentially perform dimensionality reduction using MDS from a 500-dimensional space into a 2-dimensional space. The distances between activation patterns, that MDS operates on, are defined by the representational dissimilarity matrix calculated during the process of representational similarity analysis.

3.10 Different notions of distance

This section presents the description of all the different notions of distance evaluated in this thesis. Almost all of the distance measures presented here are defined between two vectors $\mathbf{u}$ and $\mathbf{v}$. These two vectors represent the activity patterns in our fMRI data, elicited by different stimulus categories. The vectors are of length n, which is the number of voxels in the particular region of interest the vectors were obtained from. The individual components of the vectors u_i and v_i are the voxels intensity values. The descriptions were obtained from [Cha07] and are presented in alphabetical order.

3.10.1 Bray-Curtis distance

$$\frac{\sum_{i=1}^{n} |u_i - v_i|}{\sum_{i=1}^{n} (u_i + v_i)} \tag{3}$$

3.10.2 Canberra distance

$$\sum_{i=1}^{n} \frac{|u_i - v_i|}{u_i + v_i} \tag{4}$$

3.10.3 Chebyshev distance

$$max_i |u_i - v_i| \tag{5}$$

3.10.4 Cityblock distance

$$\sum_{i=1}^{n} |u_i - v_i| \tag{6}$$

3.10.5 Correlation distance

$$\frac{\sum_{i=1}^{n} (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^{n} (u_i - \bar{u})^2 \sum_{i=1}^{n} (v_i - \bar{v})^2}} \tag{7}$$

Where $\bar{u}$ and $\bar{v}$ are the means of vectors $\mathbf{u}$ and $\mathbf{v}$ respectively.

3.10.6 Cosine distance

$$\frac{\mathbf{u} \cdot \mathbf{v}}{\sqrt{\sum_{i=1}^{n} u_i^2} \sqrt{\sum_{i=1}^{n} v_i^2}} \tag{8}$$

Where $\mathbf{u} \cdot \mathbf{v}$ is the dot product between vectors $\mathbf{u}$ and $\mathbf{v}$.

3.10.7 Euclidean distance

$$\sqrt{\sum_{i=1}^{n} (u_i - v_i)^2} \tag{9}$$

3.10.8 Hamming distance

The Hamming distance between two vectors is defined as the proportion of components that differ between the vectors:

$$\frac{\sum_{i=1}^{n} I(u_i, v_i)}{n} \tag{10}$$

where I is a function returning 0 if $u_i = v_i$ and 1 otherwise.

3.10.9 Mahalanobis distance

While most of the metrics presented in this thesis measure the similarity between two vectors, the Mahalanobis distance measures the similarity between two groups of objects [DMJRM00]. In our case this means the overall similarity between all the representations of stimuli from different categories (not between two single instances of activity patterns representing stimuli from different categories). We will call this concept similarity between categories.

For starters let us define the mean elements for two different stimulus categories as $\bar{\mathbf{u}}$ and $\bar{\mathbf{v}}$. These are obtained by averaging all the individual activity patterns for a particular stimulus category component-wize. Essentially $\bar{\mathbf{u}}$ and $\bar{\mathbf{v}}$ represent the average activity patterns for two different stimulus categories. The Mahalanobis distance using these is defined as:

$$\sqrt{(\bar{\mathbf{u}} - \bar{\mathbf{v}})^T S^{-1} (\bar{\mathbf{u}} - \bar{\mathbf{v}})} \tag{11}$$

Here S^{-1} is the inverse of the pooled covariance matrix for the two categories:

$$S = \frac{cov(\mathbf{U}) + cov(\mathbf{V})}{2} \tag{12}$$

$\mathbf{U}$ and $\mathbf{V}$ are matrices containing all the activity patterns from the two different stimulus categories under consideration.

3.10.10 Kendall's tau

This metric is used for the comparison of two representational dissimilarity matrices in this thesis. It essentially describes the correlation (or more specifically agreement of the ordering of values) between the two RDMs, but in a way that is different from Spearman correlation for example. It was chosen on the basis that it was the suggested

measure for comparing RDMs in a recent review article on representational similarty analysis [NWW$^+$14b]. In this thesis we use the implementation of Kendall's τ provided by the SciPy (Open source scientific tools for Python) package [JOP14]. Their manual describes Kendall's τ between two rankings $\mathbf{X}$ and $\mathbf{Y}$ (the RDMs in our case) as:

$$\frac{P - Q}{\sqrt{(P + Q + T)(P + Q + U)}} \tag{13}$$

where P is the number of concordant pairs, Q the number of discordant pairs, T the number of ties only in $\mathbf{X}$, and U the number of ties only in $\mathbf{Y}$. If a tie occurs for the same pair in both $\mathbf{X}$ and $\mathbf{Y}$, it is not added to either T or U.

4 Results

This chapter summarizes all the results from the research carried out for this thesis. It starts by presenting the results from classification analysis described in Section 3.6, followed by a thorough example of the representational analysis pipeline used throughout this chapter. Then results from the comparison of different distance notions described in 3.10 are shown, together with their interpretation. Finally we finish off this chapter by introducing a novel use case for RSA as a technique for exploratory data analysis.

4.1 Validating the region of interest

Since the results of representational similarity analysis for a specific region of interest in the brain are not always straightforward to interpret, it would good to have a method for validating the ROI beforehand. By validation we want to determine if we can expect meaningful results from a ROI in the first place.

Since we would like to know how are objects belonging to different categories represented in different parts of the brain, the obvious starting point would be to check how well we can separate the samples in our dataset with respect to the categories.

To verify the fact that our ROI does indeed contain the necessary information to distinguish between different stimulus categories in the experiment, we trained a linear SVM classifier using data from only this region. The training procedure used fMRI activation patterns from all the experiment runs of a single subject and also included an N-fold cross-validation procedure (for a detailed description see Section 3.6). As a result of this analysis we obtained an average classification accuracy of **0.81** across all subjects. This was calculated by averaging the results across all the folds for a single subject. Clearly the accuracy is well above chance level for eight different categories (0.125) and shows that our region of interest definitely contains information about the categories of different stimuli presented to the subjects.

Figure 11 shows the classifier accuracy for a single subject represented as a confusion matrix. The confusion matrix helps us visualize classification accuracy for the different activation patterns by category and helps us identify the categories where instances are mislabeled more often. The strong diagonal in Figure 11 is an indication of a good overall classification accuracy. The categories with the highest identification accuracy are faces, houses, scrambled pictures, cats and chairs while small objects like bottles, scissors and shoes are mislabeled more often. Our results are in direct correlation with the results obtained by Haxby et al. [HGF+01] for the same dataset.
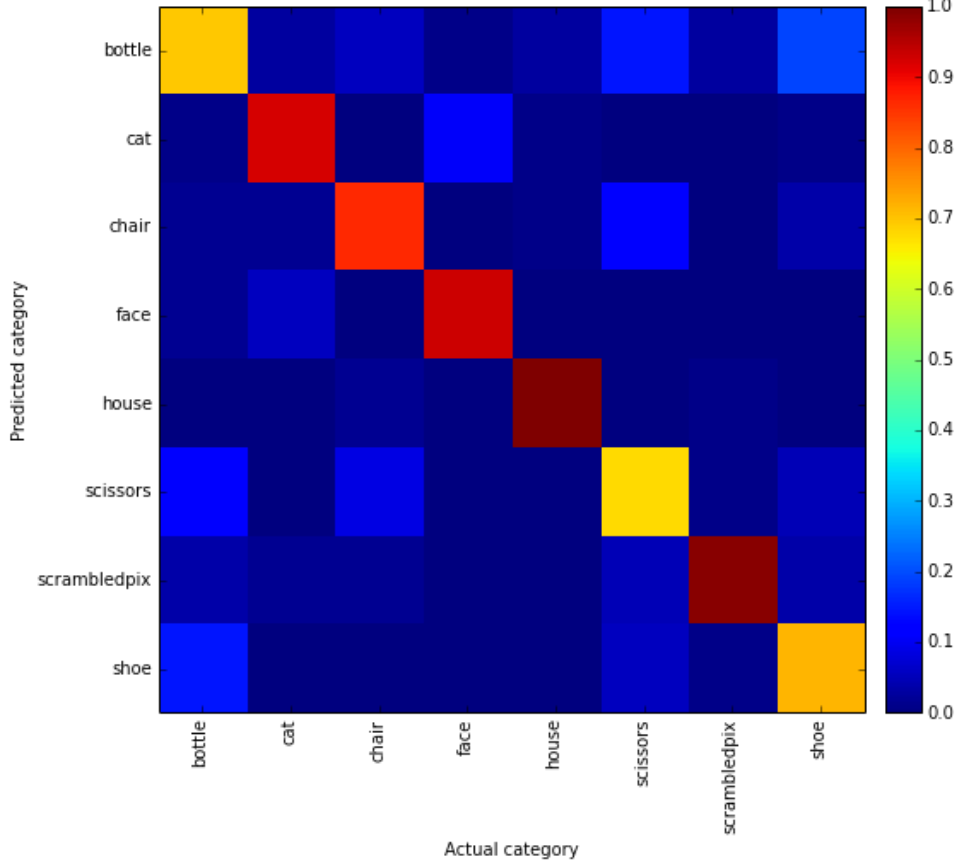
Figure 11: Confusion matrix of classifier results for subject 1. The color represents the proportion of instances for each category.

4.2 Classical representational similarity analysis

After validating our region of interest for the necessary information, we proceed to carry out classical representational similarity analysis on data from this brain region. All the steps implemented here are described in full detail in Section 3.7 together with reasoning behind each step. In this section we only describe the inputs for each step and make some remarks about the results. The analysis implemented here will serve as a template for the following sections where we run the same analysis, but with different notions of distance.

For step one we estimated the activity patterns that represent our experimental conditions, each pattern represents one of the eight categories for a stimulus image. Altogether we have 96 of these activity patterns each consisting of 577 voxels.

Next we calculate the Euclidean distance between each pair of activity patterns and assemble the results into a representational dissimilarity matrix, shown on Figure 12. The matrix is symmetrical around a diagonal of zeroes (the distance between each activity pattern and itself is zero) and is ordered by stimulus category, grouping together activity patterns elicited by stimuli belonging to the same category.

Visualizing the RDM as in Figure 12 is not very informative by itself. We can witness some distinctive square patterns formed mostly around the diagonal for some categories (indicating a close similarity between all the activity patterns comprising these particular categories) and we can also observe that some activity patterns are very dissimilar to all
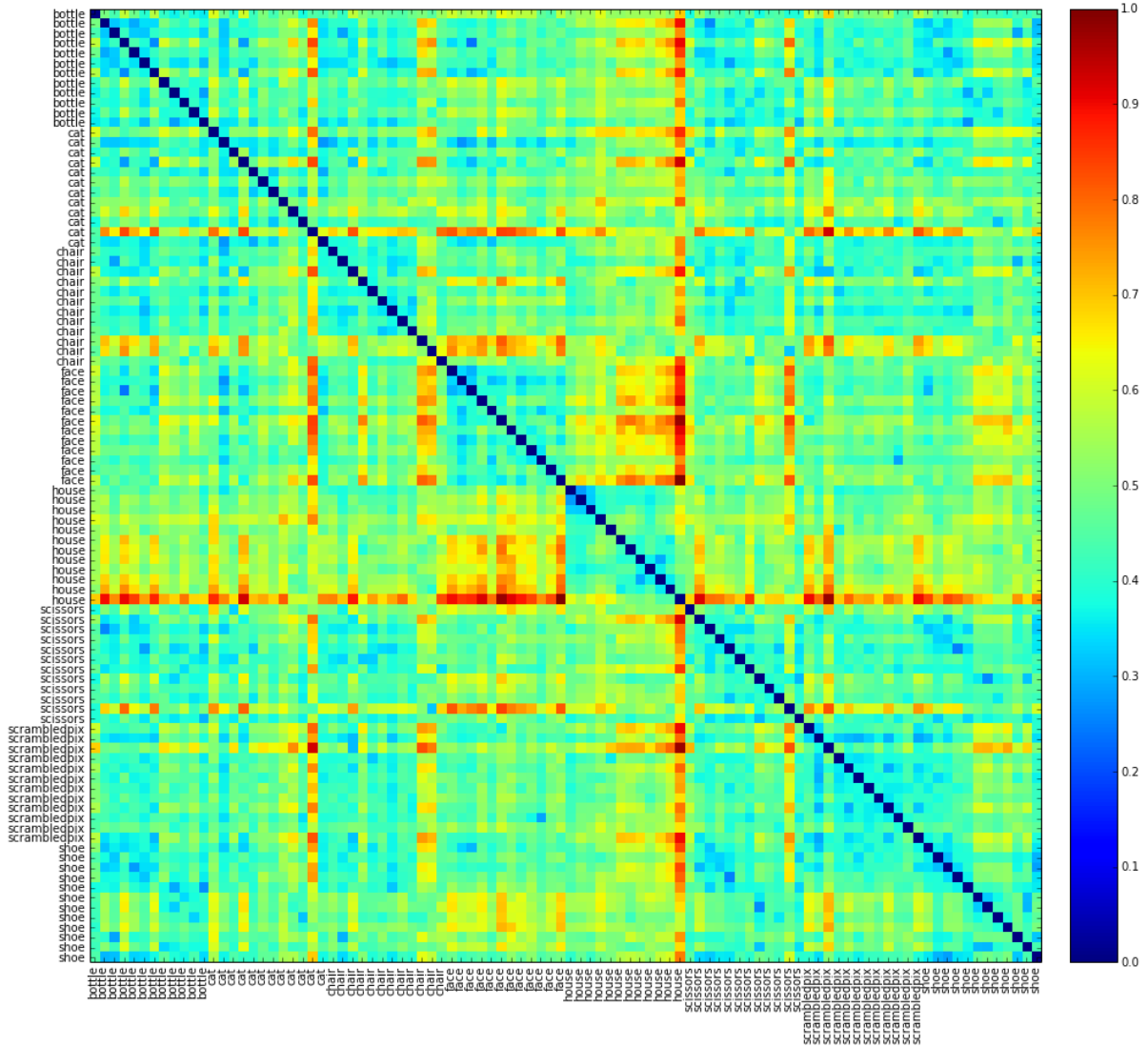
30

Figure 12: 96 x 96 Representational dissimilarity matrix. Represents the similarity structure between each pair of 96 different activity patterns (grouped by the stimulus category that ellicited them). Distance values are scaled to be between 0 and 1, the latter representing the maximum distance calculated between any two representations.

other activity patterns not belonging to their particular category. Still this does not give us a clear way to visualize the relation of representations of different stimulus categories.

To try to visualize the relationship between representations of different stimulus categories more clearly, we reordered the rows and columns of the representational dissimilarity matrix. The new ordering was generated by clustering the distances of activity patterns in the RDM using hierarchical clustering. This process is described in more detail in Section3.8. The output of hierarchical clustering is a dendrogram where the leaf nodes are our activity patterns and they are ordered based on their similarity. We used this ordering for the rows and columns of the dissimilarity matrix. The main idea for this sort of reordering of the RDM is to visualize "natural" clusters of activity patterns in the dataset. The results are shown on Figure 13.
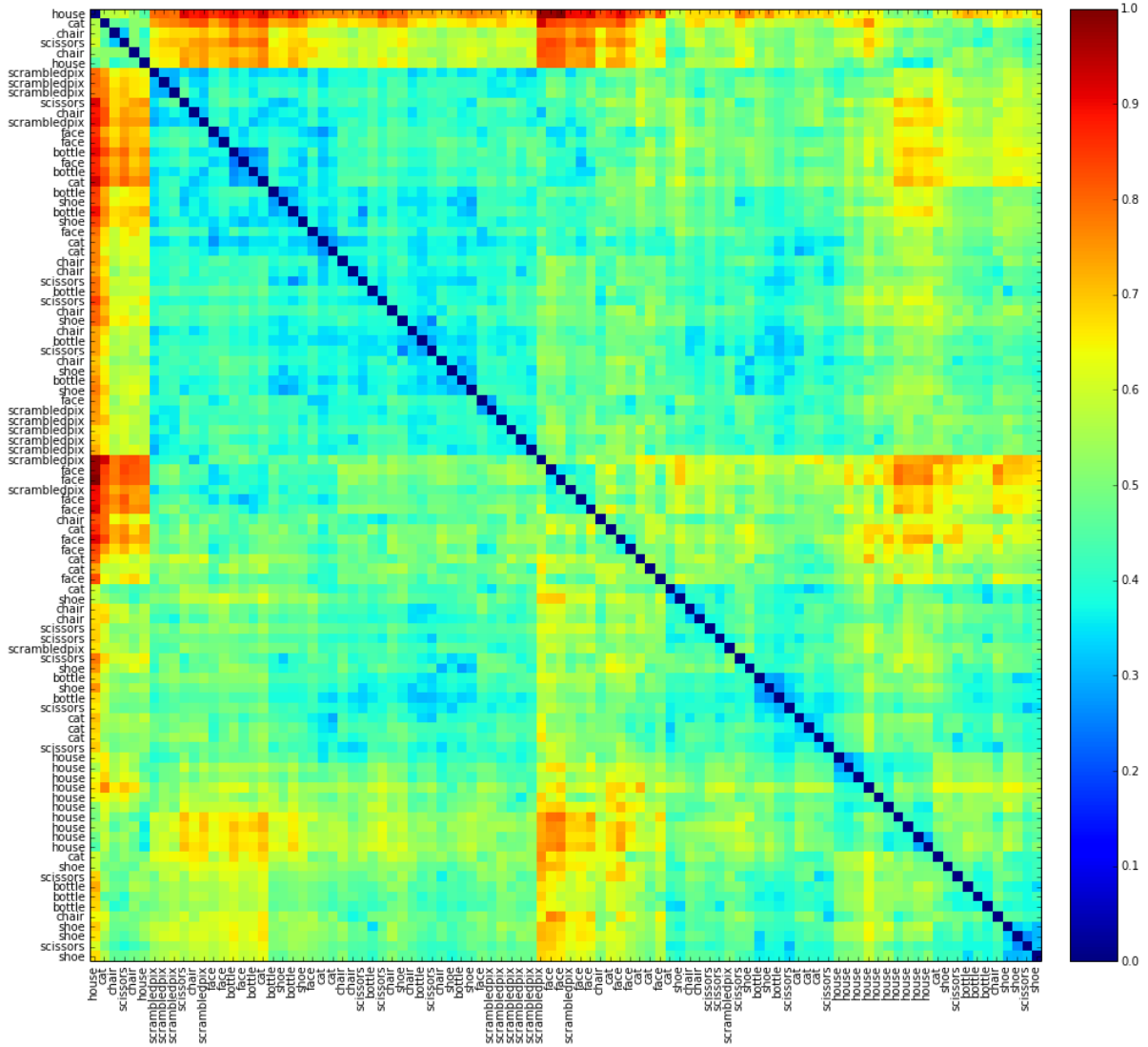

Figure 13: Representational dissimilarity matrix with rows and columns reordered to place similar activity patterns together.

In this new matrix we see more distinctive rectangular patterns and not all of them are around the diagonal anymore either. This is a clear indication that our dataset contains some structure with respect to our chosen distance metric (the Euclidean distance). We would expect that representations of activity patterns caused by the same stimulus

category would be close together, but as we see from Figure 13 this is not always the case. Although there is some clustering in that respect, like houses, the overall similarity structure seems to be more complex.

The third step in our analysis involves the creation of a model. For this we chose a conceptual model inspired by one described in [KMB08]. Our model describes a hypothetical region of the brain where representations of animate and inanimate objects are very dissimilar. If we presented our model with the exact same stimuli as were presented to the subjects in the experiment and carried out the first two steps of RSA, the resulting RDM from the model would look like the one depicted on Figure 14.
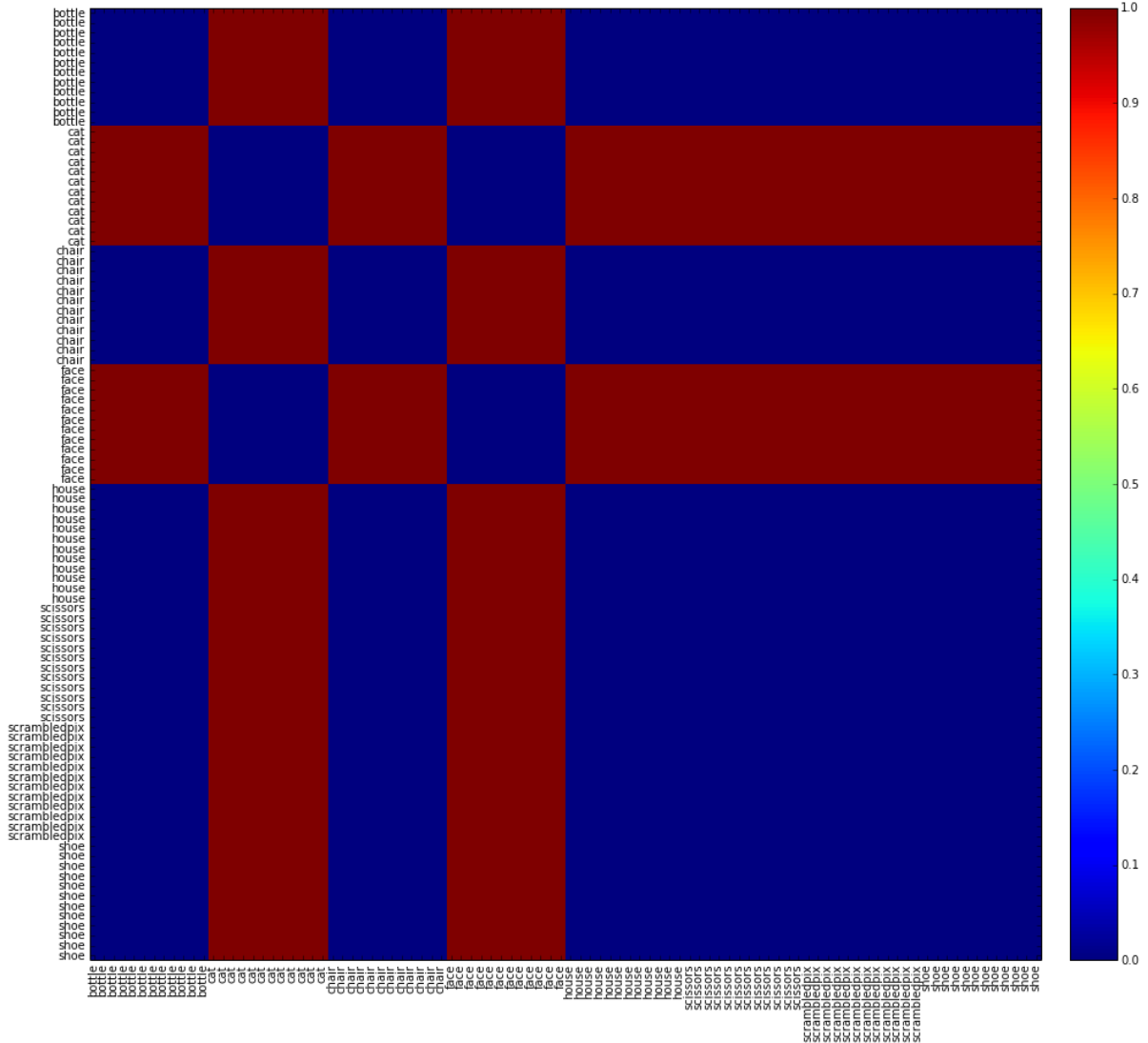

Figure 14: Representational dissimilarity matrix from a model distinguishing perfectly between animate and inanimate objects.

The RDM from the model depicts representations of stimulus categories of animate objects (cats and faces) being maximally similar to each other while at the same time maximally dissimilar to representations of inanimate objects (bottles, chairs, houses, scissors, scrambled pictures and shoes). This model is of course very simplistic as it does not model any noise or other variables except our conceptual notion of animate and inanimate objects being very distinguishable. In fact we do not have any reasonable

neuroscientific justification for believing that our experimental region of interest (the ventral temporal cortex) is able to distinguish between animate and inanimate objects at all. Still for the purposes of this thesis, the model will serve well for comparison with RDMs based on actual activity patterns as our goal is not to prove the validity of this model for the experimental region of interest.

The Kendall's τ coefficient between the RDM on Figure 12 and our model is **0.16**. The interpretation behind this is that the correlation between representations of cognitive states within the model and actual measured activity patterns is very weak. This implies that the actual information processing occurring in the ROI interest under study is very different from the assumptions of our simplified model (as was expected) and that this particular area in the brain does not make a very clearly identifiable distinction between animate and inanimate objects.

For the final step we visualize the similarity structure of the representations of cognitive states contained in the RDMs by multidimensional scaling. Figure 15 shows two plots, both based on the RDM in Figure 12. The distances of individual points in the plots represent the actual distances of the different neural representations (voxel activation patterns) for each experiment block (as measured by Euclidean distance). The motivation behinds using two plots instead of one, is to both visualize the variance in similarity between representations of the same stimulus category and the distance between representations of different categories more clearly at the same time.
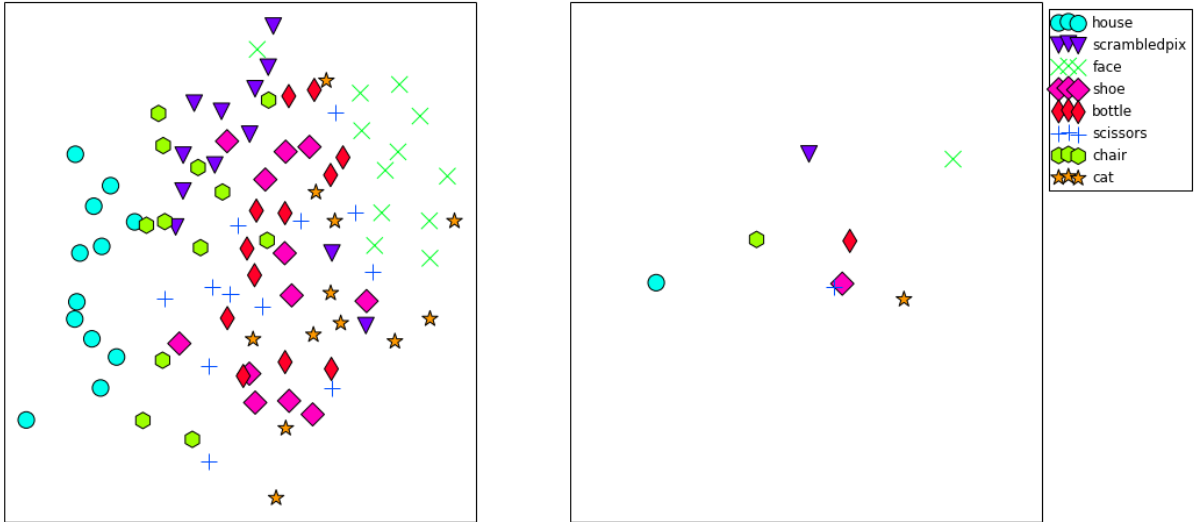

Figure 15: Multidimensional scaling of the representations of cognitive states ellicited by the stimulus images. On the left: each dot is the neural representation of a stimulus block in the experiment (the representation of the category of a stimulus). On the right: centroids for the individual points in the left plot averaged by category.

From Figure 15 we can gather that representations of houses and faces seem to exhibit a relative high within category similarity between instances (tighter grouping), while the representations of smaller objects like shoes, bottles and scissors have a not so clearly defined similarity structure and are therefore more difficult to distinguish, as individual representations from these categories are intermingled with each other. Although the results are not directly comparable, intuitively we could argue that they coincide somewhat with the results of the classifier analysis implemented in Section 4.1, where small objects were mislabeled more often, while houses and faces had a much higher accuracy.

For visual comparisons we also ran MDS on our model that distinguishes only between animate and inanimate objects, the results are shown on Figure 16.
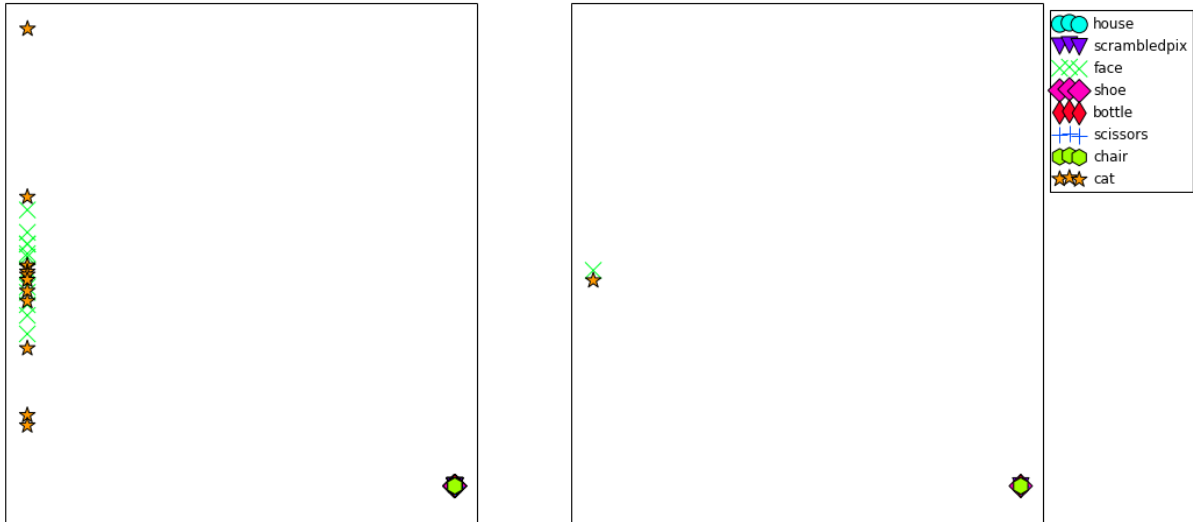

Figure 16: MDS results from a conceptual model distinguishing perfectly between animate and inanimate objects.

Purely by visual comparison we are able to draw additional verification for the low Kendall coefficient calculated earlier. As in the model we see the centroids for the faces and cats category very close together and at the same time very far apart from all the rest of the categories, while in the plots of the actual representations, we cannot really say that cats are more closer to faces than for example bottles or shoes.

4.3 The effect of distance on RSA

The main goal of this thesis is to study different notions of distance between the activity patterns measured in the subjects and the effect they have in the context of representational similarity analysis. This section summarizes the results we obtained by comparing 9 different distance notions (described in Section 3.10).

In order to compare the effects of different distance functions, we first carried out RSA as described in Section 4.2 substituting the Euclidean distance used in that section with all the different notions of distance in turn. The results from multidimensional scaling for each different notion of distance are shown on Figures 17 - 24, visualizing for each distance the similarity structure between representations of all the different stimulus blocks on the left and the similarity between stimulus categories on the right. The right plots visualize the centroids of individual points on the left by category.

The multidimensional scaling plots are useful to visualize general tendencies in the data and how stable these tendencies are in the face of changing distance functions. As the whole idea of RSA is to abstract away from individual activation patterns, we would like to know if the distance functions used to estimate the representational similarity structures of activity patterns have any effect on this higher level of abstraction. For this we compared the RDMs constructed with different distance functions to the RDM of our animate-inanimate model via the Kendall's τ coefficient (Table 1).
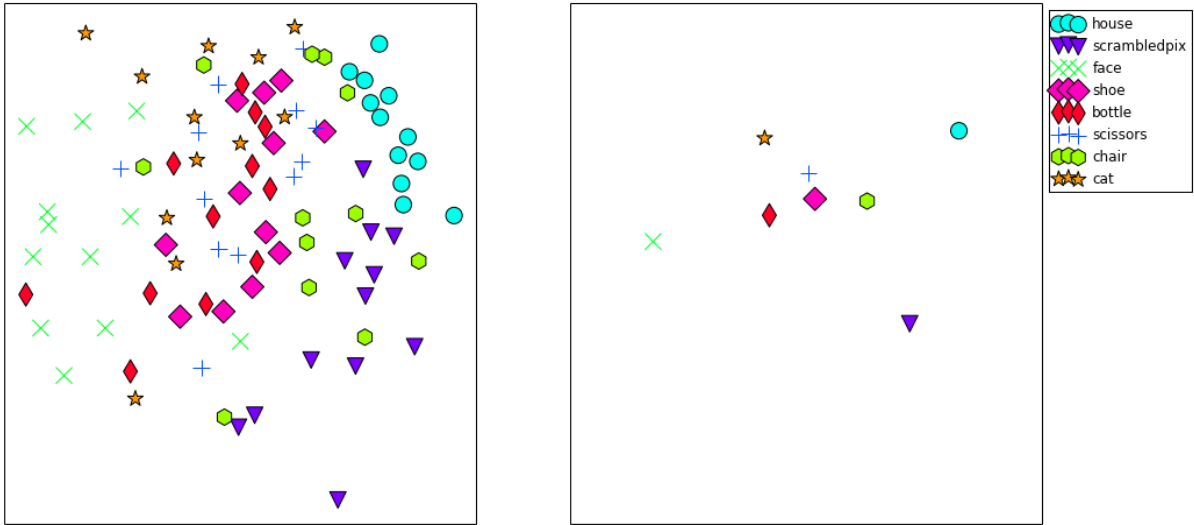

Figure 17: MDS results using Bray-Curtis distance

Some distinctive patterns emerge from the plots on Figures 17 - 24. For almost all different distance notions, representations from the "face" and "house" categories exhibit the greatest within category similarity (tight grouping on most of the plots). Representations from all the other categories do not exhibit such a clearly defined similarity structure within their category. Representations of faces and houses are also consistently most dissimilar to each other as indicated by the centroid plots on the right. Some distance notions like Chebyshev and Correlation distance seem to exhibit a tighter grouping also in the "cat" category, but the results are far from conclusive.

Both the MDS plots and model correlation coefficients on Table 1 show Hamming distance as the odd ball out. Recall from Section 3.10 that Hamming distance between two vectors represents the proportion of those vector elements between the two n-dimensional
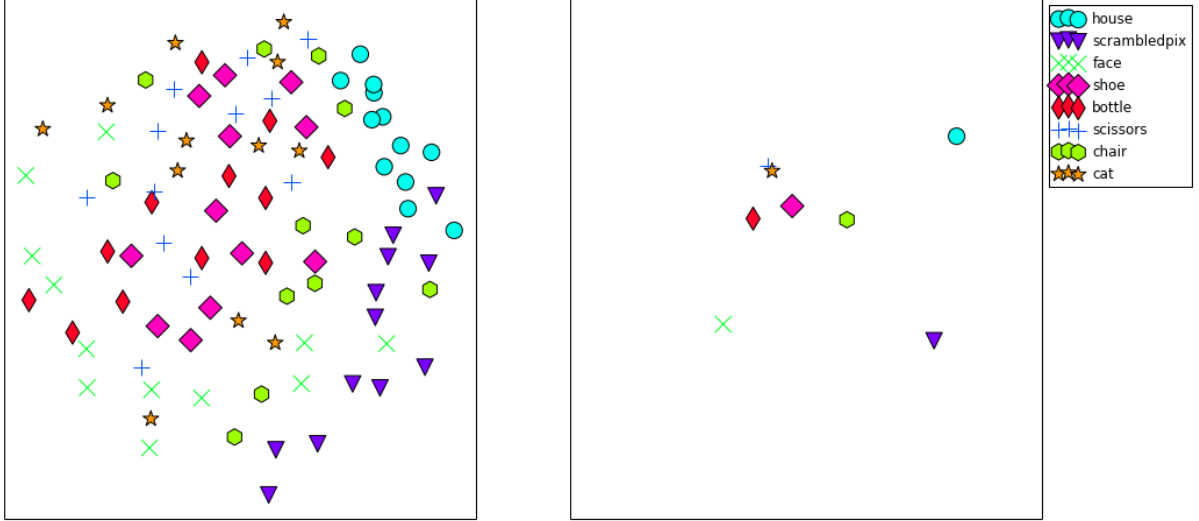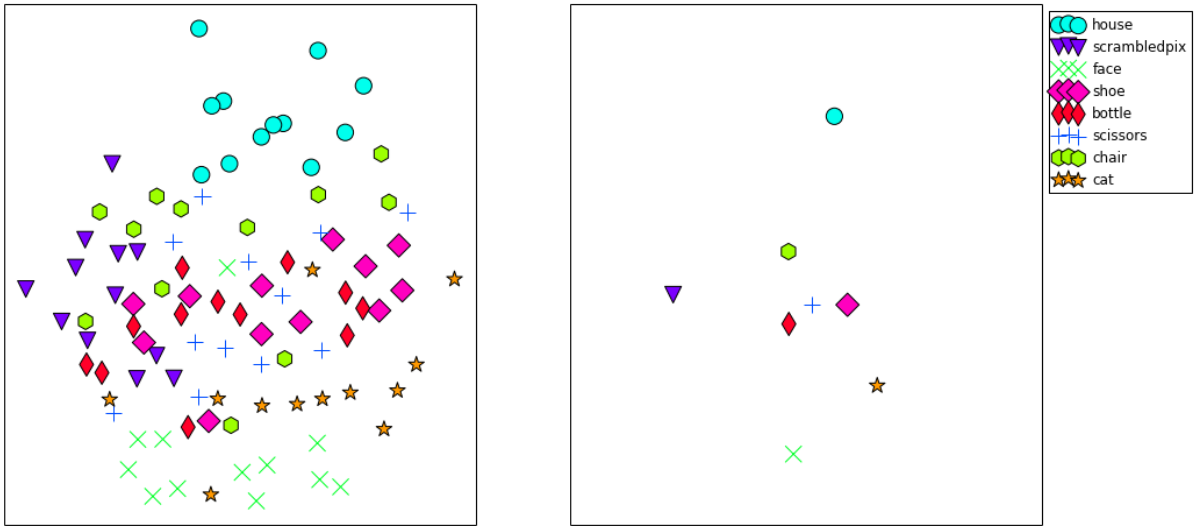
Figure 18: MDS results using Canberra distance

Figure 19: MDS results using Chebyshev distance

Table 1: Comparison of activity pattern RDMs to the model RDM

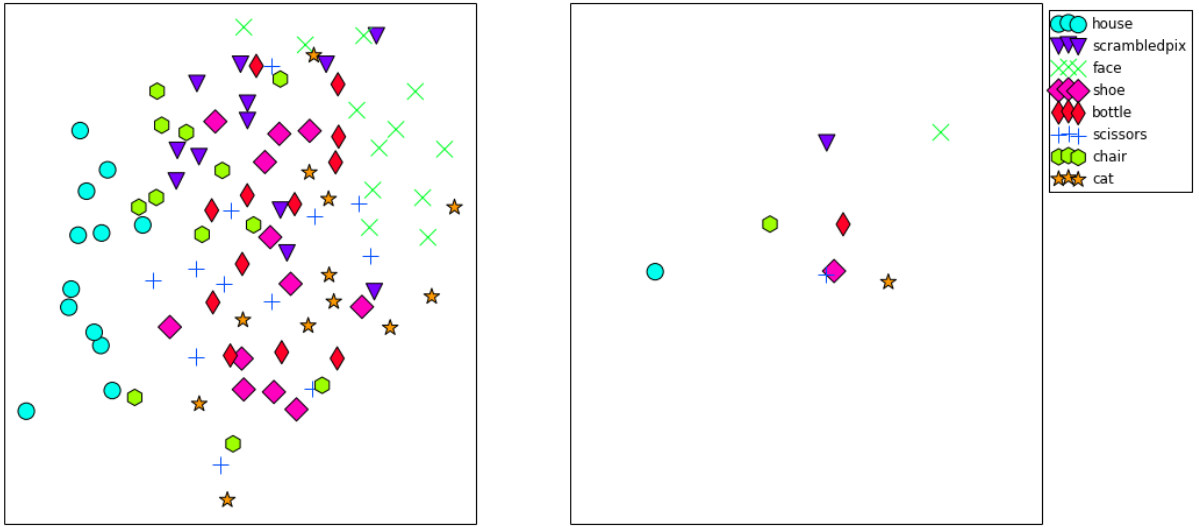| Distance notion	Kendall's τ coefficient
Bray-Curtis	0.22
Canberra	0.21
Chebyshev	0.18
Cityblock	0.16
Correlation	0.22
Cosine	0.25
Euclidean	0.17
Hamming	0.08
Mahalanobis	0.34

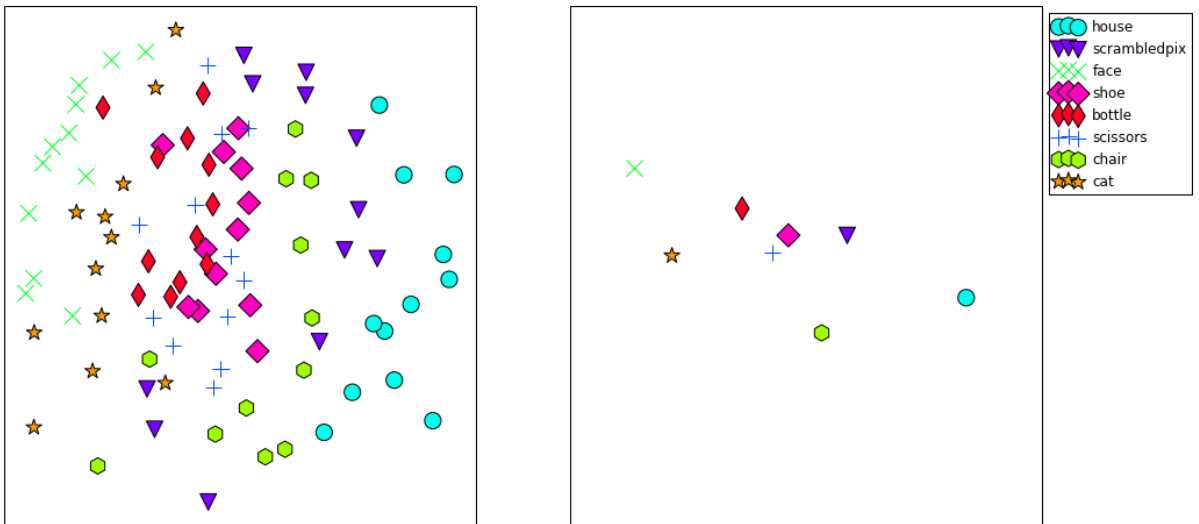Figure 20: MDS results using Cityblock distance
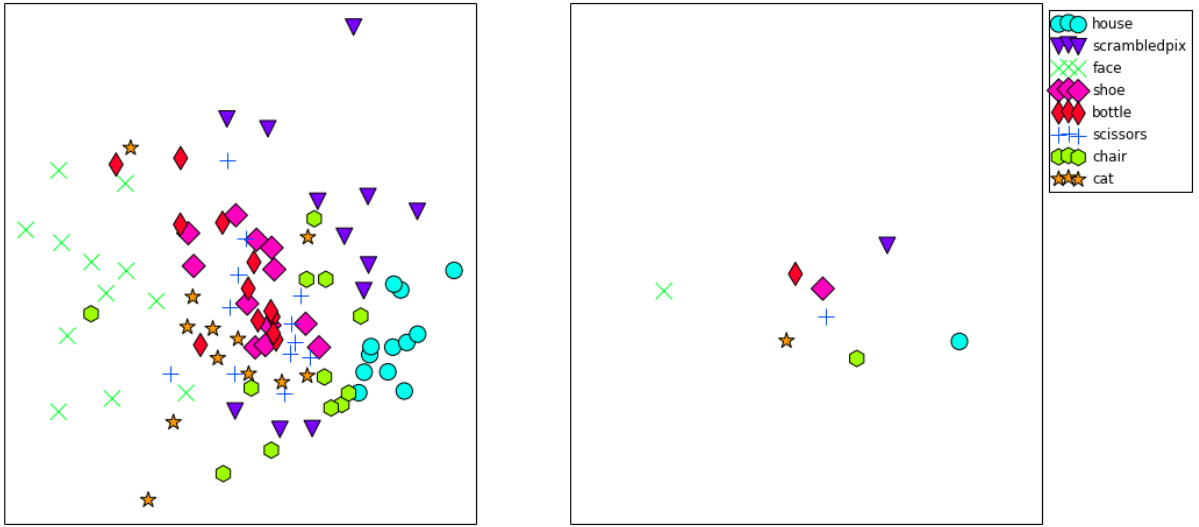

Figure 21: MDS results using Correlation distance
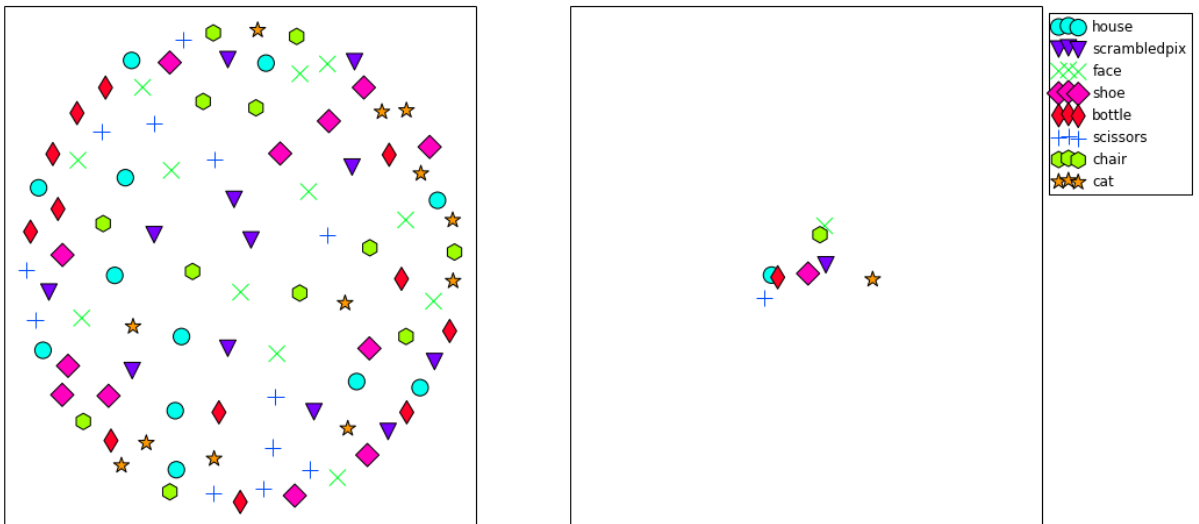
Figure 22: MDS results using cosine distance

Figure 23: MDS results using Hamming distance

vectors **u** and **v** which disagree. As such it is only meant to be used for vectors of discrete elements. Since our fMRI image vectors contain continuous values, it really makes no sense whatsoever in using the Hamming distance to assess the similarity between them. In fact the only reason Hamming distance was added to this comparison in the first place, was to serve as an example that not all distance metrics between any two vectors are appropriate in the domain of fMRI data analysis. Adding such a nonsensical metric in this context can also serve as a sort of additional validation of the analysis pipeline in the sense of determining how easily we can detect the uselessness of this metric from the results.

As for the comparison with the model, when we look at Table 1 (excluding the Hamming distance), indeed there exist some differences in the correlation between the different RDMs and the model RDM. Especially representations estimated using the Mahalanobis distance do seem to correlate more with the model than representations from any other distance notion. This might have something to do with the way we generated the RDM for Mahalanobis distance. Since the Mahalanobis distance metric measures the similarity between stimulus categories not between the representations of individual stimulus blocks as all the other metrics, we had to interpolate the RDM for Mahalanobis distance in order to compare it with the model. As the model RDM is a 96 x 96 matrix and the initial result from Mahalanobis was a 8 x 8 matrix (similarity between each of the 8 categories), we used nearest neighbor interpolation to construct a 96 x 96 matrix from the original 8 x 8 matrix. This essentially means that all the individual representations in the interpolated matrix for a single category are identical and represent the average of all the individual samples in some sense.

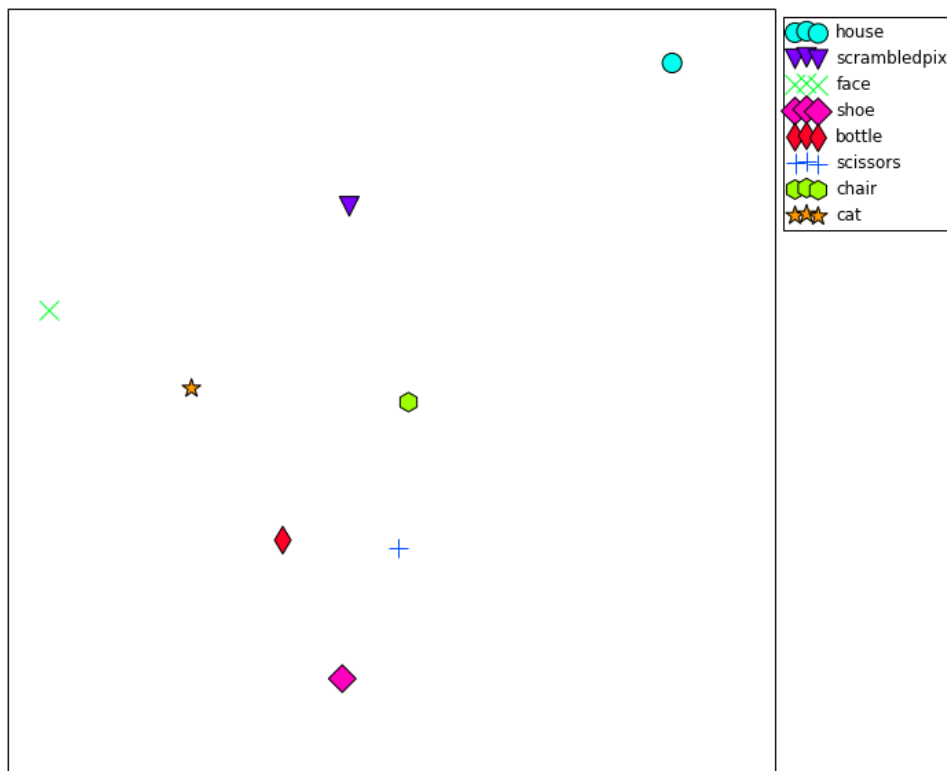For a better comparison of the results from multidimensional scaling, we calculated the

Figure 24: MDS results using Mahalanobis distance

pairwise distances of all the centroids from the right plots and ordered them by distance from smallest to largest. Since in essence the distances between centroids represent the average similarity between stimuli from different categories, we wanted to see how this ranking of pairwise distances changed with respect to the distance notion used. This difference in the rankings is visualized on Figure 25. Since this visualization is just another technique for representing the results from all the MDS plots in a more concise way, we can draw the exact same conclusions as before. The overall tendencies for all the different distance notions are largely the same - the distance between faces and houses is the greatest with all the meaningful distance functions used, while there are also considerable similarities in the pairwise distances for the small objects. Nevertheless there are specific differences, for example the distance between scrambled pictures and faces is estimated very differently between the Euclidean distance (13th in the ranking) and a group consisting of cosine, Canberra and Bray-Curtis distances (second largest distance).
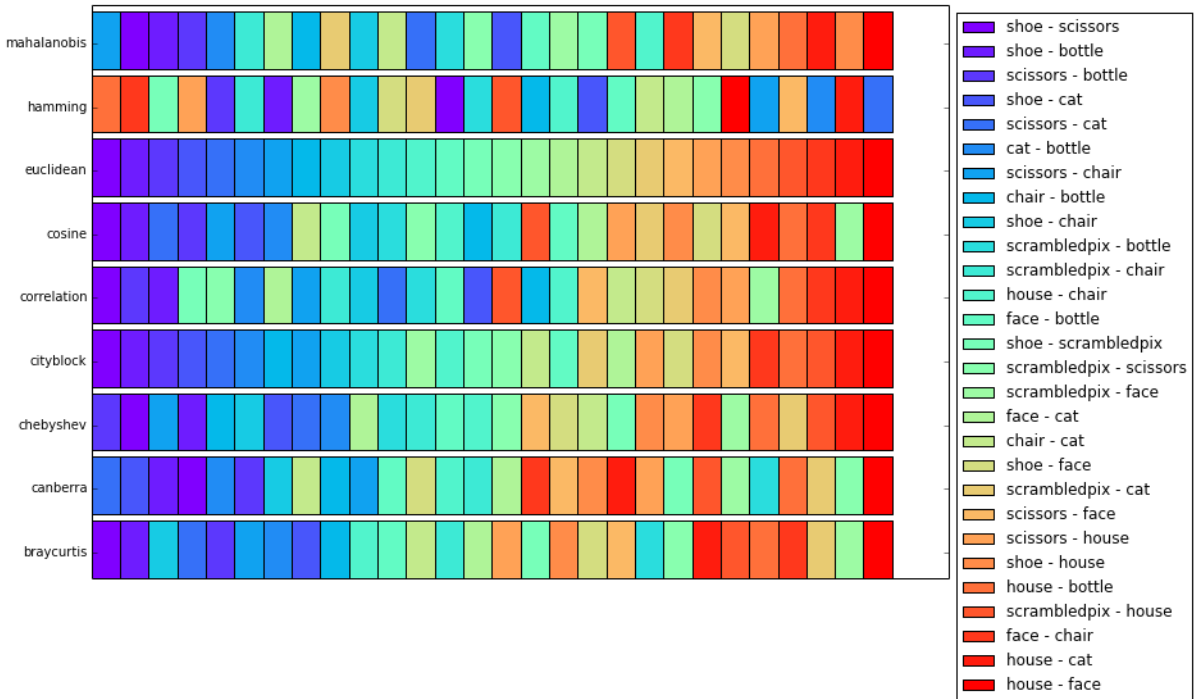

Figure 25: The ranking of pairwise distances between average representations of different stimulus categories for each notion of distance used. The pairs of categories are ordered and colored according to their ranking in the Euclidean distance, from smallest pairwise distance (purple) to largest pairwise distance (red). This allows us to visualize the differences of ranking, by taking Euclidean distance as a reference

For a final comparison of the different distance functions, we calculated the second order similarity between their results (the similarity between each pair of RDMs which in turn represent the similarity structures estimated from the underlying activity patterns with different notions of distance). Figure 26 depicts the relative similarity of the different distance notions under consideration.
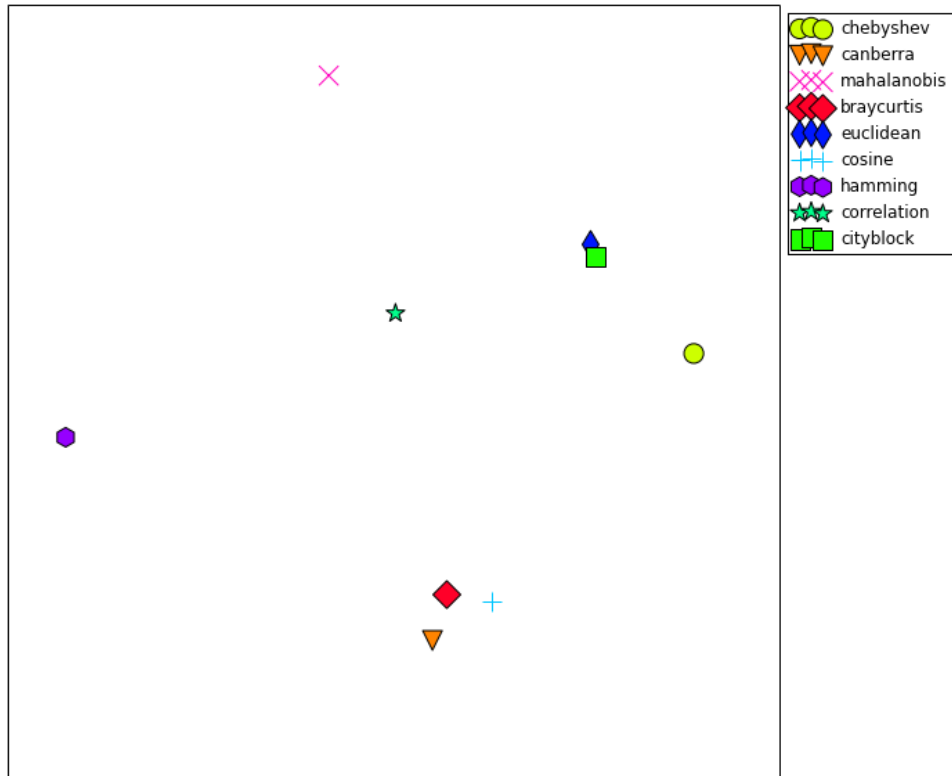
Figure 26: MDS visualization of second order similarity between all the RDMs calculated with different distance notions. $1 - Kendall's\tau$ is used to calculate the dissimilarity between RDMs.

4.4   RSA as a tool for assessing data quality

While doing research for this thesis and analysis tests, we ended up visualizing many different configurations of representational distance matrices and multidimensional scaling plots. This constant visual feedback led us to discover a bug in our data preprocessing pipeline which would have been very difficult to detect by other means. Therefore we propose the notion that several steps from the representational similarity analysis pipeline could also be used for assessing the effects of different data preprocessing techniques and other transformations.

Our dataset is a timeseries of fMRI volumes each consisting of hundreds of voxels recorded at different timepoints. When dealing with multivariate data such as this, it can be very difficult to visualize what effect a particular transformation (like detrending) has on the dataset as a whole. Possibilities include visualizing raw voxel intensities, which is not very informative. Another sometimes useful technique is to visualize the timeseries of one or a couple of individual voxels, but still this is lacking a global point of view. Combine these problems with the inherently low signal to noise ratio of fMRI data and the situation looks pretty bleak.

It turns out that we can actually leverage RSA to help alleviate the lack of globally effective visualization options. Since RSA handles the individual data samples (fMRI volumes) on a more abstract level, we are able to observe changes caused by transformations of the original samples in a more meaningful way. As an example consider the two represetational dissimilarity matrices on Figure 27.
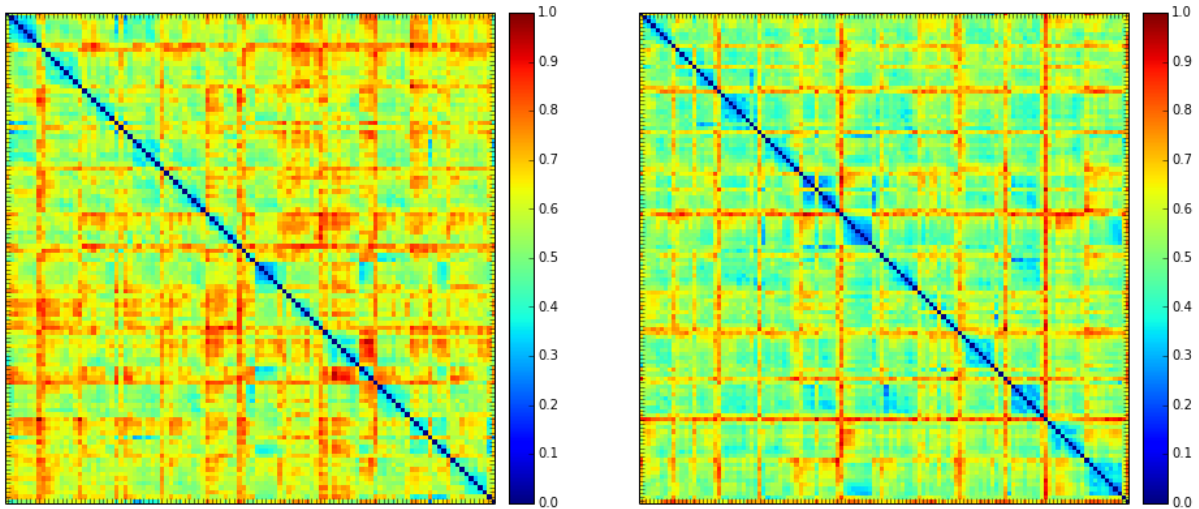

Figure 27: Two RDMs representing the similarity structure between individual fMRI volumes of the "faces" category.

They both represent the similarity structure of individual fMRI volumes from only the "faces" category. Our underlying assumption here is that since the representations are caused by the same stimulus category, they should be pretty similar to each other. The distinguishing block texture we can observe on them stems from the fact that samples are recorded on different experiment runs and recordings from on different runs are bound to be more dissimilar. Still we notice that even samples from the same run show up as being very dissimilar in the left RDM. This led us to believe that some of the samples in the dataset vere mislabeled, which upon the revising of our preprocessing pipeline turned

out to be correct. The right RDM represents the same data, but with correct processing of the labels of individual fMRI volumes.

This example also illustrates the fact that it can be beneficial at times to visualize only slices from the entire RDM to focus on specific questions related to only some of the conditions. Visualizing the entire RDM can be useful for detecting areas that can be interesting for more detailed study.

Altough we did validate our data with a classification analysis (Section 4.1), even with the mislabeled samples, the average classification accuracy was way beyond chance level. The implication here are that by relying on the accuracy metric alone, this discovery would never have been made. With the corrected pipeline, we managed to improve classification accuracy by another 12%.

The above is just one example of how RSA can be leveraged to provide a view into the data from a different angle. This is somewhat analogous to how a Fourier transform is used in signal processing to convert a time domain signal into frequency domain, in order to study different aspects of the same signal [Bra65]. We think that RSA can also be used to visualize the effects of other preprocessing steps like detrending, normalization and z-scoring.

5 Discussion

As far as we know, we are the first to conduct a systematic study into the effects of different distance notions in the context of representational similarity analysis. Most literature about RSA seems to prefer the use of Correlation distance to assess the similarity between activation patterns ([KMB08], [KMR⁺08]), because it normalizes for both the mean level of activity and the variability in the activation patterns. Although there is also mentioning of Euclidean and Mahalanobis distances in the context of RSA ([NWW⁺14a]), we have not seen any thorough comparisons between them.

As we evaluated 9 different notions of distance, we did so with no particular neuro-scientific research question to optimize for. In other words, we were not trying to find the "best" notion of distance, as the utility of a particular distance measure would only become apparent in the context of some specific research goals. An example of this would be a study where in addition to the neural representations, the reaction times of a subject for experimental stimuli would also be recorded. In this case RSA could be used to relate the neural representations of specific stimuli to the reaction times corresponding to those same stimuli. Here a useful metric would be one that maximises the correlation between the neural representations and the reaction times. Since we did not have a dataset with such behavioral data at hand, we evaluated the distance measures against a conceptual model instead.

We showed that although there are differences in the final results of RSA depending on the distance measure used to estimate the similarity between initial activity patterns, the differences in our case did not change the most prevalent trends in the results - all the distance measures exhibited noticeable grouping for the representations of houses and faces. What is more interesting however is that there exist considerable similarities between some distance notions. For example the Bray-Curtis, Canberra and Cosine measures produce different but very similar pairwise distance rankings of the representations of stimuli. The same can be said for Euclidean and Cityblock distances, which are even more similar. Being aware of this kind of grouping among the measures can save researchers some time when they are evaluating different distance notions for a particular use case. They can then only focus on measures that are from different groups. Based on our results the optimal set of distances to evaluate for a research problem would contain Mahalanobis, Correlation, Euclidean, Chebyshev and Cosine metrics.

This kind of grouping of measures demonstrated by our analysis is not surprising however. From a mathematical perspective, the notions of distance are not fully independent and one could have expected similar behavior for a few of them as indicated by the grouping. As we implemented only experimental comparisons of the distance measures, we refer the reader to a general overview paper where the semantic and syntactic similarity is also taken into account when doing comparisons [Cha07].

Next we will propose some additional ideas and variations to the techniques used in this thesis for further research in the future:

- In Section 4.4 we presented another use case for RSA as a tool to visualize the effects different transformations on the input dataset have on a global level. We postulated that RSA might be useful to assess the effects different preprocessing steps like detrending, intensity normalization or z-scoring have on the data visually. Another direction would be to try to automate the process for finding optimal preprocessing steps. We could try different ranks of polynomials for the detrending process, carry out RSA for each of them and define some metric to assess the quality

of the resulting RDMs (for example the average similarity between representations of the same category). This is a tricky proposition, as we know now that the results would depend not only on the metric defined for the RDMs, but also from the notion of distance used to estimate the similarity between activity patterns themselves.

- The list of notions evaluated for this thesis is by no means comprehensive. We tried to select a good subset of known distance measures and the main goal was to show whether they have any difference at all in the context of RSA. The obvious suggestion here would be to continue the research with more distance measures. It would be especially interesting to see how modality specific distance measures compared with our results. By that we mean measures that are specifically adapted to take into account also the spacial structure of fMRI images for example. A set of such novel measures is proposed by Novatnack et al. [NCS+08].

- All analysis in this thesis was carried out on data coming only from a single subject. Since the dataset contained the neural responses for six different subjects, comparisons could be made by carrying out the same analysis for different subjects.

- We described the concept of defining a region of interest in the brain functionally in Section 3.5. One possible area of interest would be to use some functionally defined regions in determining if there are some regions where the overall similarities between distances as illustrated on Figure 26 differ significantly.

- Multidimensional scaling is extensively used in this thesis to visualize similarity structures contained in representational similarity matrices. In a review article on RSA [NWW+14b] another method besides MDS is also suggested for the same purpose, called t-SNE [VdMH08]. We actually evaluated t-SNE for visualizing the distances between representations in this thesis. In some cases it would yield more aesthetically pleasing results than MDS, but that required careful tuning of the input parameters. MDS was chosen over t-SNE mainly because the results we obtained with the latter were inconclusive. Still t-SNE could prove to be a viable alternative for MDS, but this needs more research.

6   Conclusions

In this thesis we presented a thorough overview of how to implement representational similarity analysis on fMRI data using a well known dataset in neuroscience as the basis. We demonstrated all the steps in a typical RSA pipeline starting from the preprocessing of data and ending with the interpretation of the results. We also showed how the visualizations of representational dissimilarity matrices can be improved by hierarchical clustering.

The main goal of this thesis was to evaluate 9 different notions of distance to be used for estimating the similarity between representations of different stimuli in the brain and the effect these distance notions have on the overall result of RSA. We concluded that the results of RSA indeed differ with respect to the distance measures used between activity patterns and that there exist considerable similarities between different groups of distance measures.

In addition to the comparison of distance metrics, we also proposed a novel use case for representational similarity analysis as a tool for visualizing the effects of different transformations on the input data on a global level.

References

[AB06]   Edson Amaro and Gareth J Barker. Study design in fmri: basic principles. *Brain and cognition*, 60(3):220–232, 2006.

[BG05]   Ingwer Borg and Patrick JF Groenen. *Modern multidimensional scaling: Theory and applications.* Springer Science & Business Media, 2005.

[Bra65]   Ron Bracewell. The fourier transform and iis applications. *New York*, 1965.

[CAB+04]   Robert W Cox, John Ashburner, Hester Breman, Kate Fissell, Christian Haselgrove, Colin J Holmes, Jack L Lancaster, David E Rex, Stephen M Smith, Jeffrey B Woodward, et al. A (sort of) new image data format standard: Nifti-1. *Human Brain Mapping*, 25:33, 2004.

[Cha07]   Sung-Hyuk Cha. Comprehensive survey on distance/similarity measures between probability density functions. *City*, 1(2):1, 2007.

[Dev07]   H Devlin. What is functional magnetic resonance imaging (fmri)? Psych Central, 2007. `http://psychcentral.com/lib/what-is-functional-magnetic-resonance-imaging-fmri/0001056`.

[DMJRM00] Roy De Maesschalck, Delphine Jouan-Rimbaud, and Désiré L Massart. The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, 50(1):1–18, 2000.

[EZB13]   Joset A Etzel, Jeffrey M Zacks, and Todd S Braver. Searchlight analysis: promise, pitfalls, and potential. *Neuroimage*, 78:261–269, 2013.

[Gre07]   Michael Greenacre. *Correspondence analysis in practice.* CRC Press, 2007.

[HDO+98]   Marti A. Hearst, Susan T Dumais, Edgar Osman, John Platt, and Bernhard Scholkopf. Support vector machines. *Intelligent Systems and their Applications, IEEE*, 13(4):18–28, 1998.

[HGF+01]   James V Haxby, M Ida Gobbini, Maura L Furey, Alumit Ishai, Jennifer L Schouten, and Pietro Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539):2425–2430, 2001.

[HHS+09]   Michael Hanke, Yaroslav O Halchenko, Per B Sederberg, Stephen José Hanson, James V Haxby, and Stefan Pollmann. Pymvpa: A python toolbox for multivariate pattern analysis of fmri data. *Neuroinformatics*, 7(1):37–53, 2009.

[HMH04]   Stephen José Hanson, Toshihiko Matsuka, and James V Haxby. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage*, 23(1):156–166, 2004.

[JOP14]   Eric Jones, Travis Oliphant, and Pearu Peterson. {SciPy}: Open source scientific tools for {Python}. 2014.

[KB07]   Nikolaus Kriegeskorte and Peter Bandettini. Analyzing for information, not activation, to exploit high-resolution fmri. *Neuroimage*, 38(4):649–662, 2007.

[KK13]   Nikolaus Kriegeskorte and Rogier A Kievit. Representational geometry: integrating cognition, computation, and the brain. *Trends in cognitive sciences*, 17(8):401–412, 2013.

[KMB08]   Nikolaus Kriegeskorte, Marieke Mur, and Peter Bandettini. Representational similarity analysis–connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2, 2008.

[KMR⁺08]   Nikolaus Kriegeskorte, Marieke Mur, Douglas A Ruff, Roozbeh Kiani, Jerzy Bodurka, Hossein Esteky, Keiji Tanaka, and Peter A Bandettini. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6):1126–1141, 2008.

[Kru64]   Joseph B Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

[Log03]   Nikos K Logothetis. The underpinnings of the bold functional magnetic resonance imaging signal. *The Journal of Neuroscience*, 23(10):3963–3971, 2003.

[NCS⁺08]   John Novatnack, Nicu Cornea, Ali Shokoufandeh, Deborah Silver, Sven Dickinson, Paul Kantor, and Bing Bai. A generalized family of fixed-radius distribution-based distance measures for content-based fmri image retrieval. *Pattern Recognition Letters*, 29(12):1726–1732, 2008.

[NPDH06]   Kenneth A Norman, Sean M Polyn, Greg J Detre, and James V Haxby. Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in cognitive sciences*, 10(9):424–430, 2006.

[NWW⁺14a] Hamed Nili, Cai Wingfield, Alexander Walther, Li Su, William Marslen-Wilson, and Nikolaus Kriegeskorte. A toolbox for representational similarity analysis. *PLoS computational biology*, 10(4):e1003553, 2014.

[NWW⁺14b] Hamed Nili, Cai Wingfield, Alexander Walther, Li Su, William Marslen-Wilson, and Nikolaus Kriegeskorte. A toolbox for representational similarity analysis. *PLoS computational biology*, 10(4):e1003553, 2014.

[OJAH05]   A O'toole, Fang Jiang, Hervé Abdi, and J Haxby. Partially distributed representations of objects and faces in ventral temporal cortex. *Cognitive Neuroscience, Journal of*, 17(4):580–590, 2005.

[PBM⁺13]   Russell A Poldrack, Deanna M Barch, Jason P Mitchell, Tor D Wager, Anthony D Wagner, Joseph T Devlin, Chad Cumba, Oluwasanmi Koyejo, and Michael P Milham. Toward open sharing of task-based fmri data: the openfmri project. *Frontiers in neuroinformatics*, 7, 2013.

[PG07]   Fernando Pérez and Brian E. Granger. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, 9(3):21–29, May 2007.

[PMB09]   Francisco Pereira, Tom Mitchell, and Matthew Botvinick. Machine learning classifiers and fmri: a tutorial overview. *Neuroimage*, 45(1):S199–S209, 2009.

[PPFR88]   Michael I Posner, Steven E Petersen, Peter T Fox, and Marcus E Raichle. Localization of cognitive operations in the human brain. *Science*, 240(4859):1627–1631, 1988.

[PVG+11]   Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12:2825–2830, 2011.

[RKB12]   C Rorden, HO Karnath, and L Bonilha. Mricron dicom to nifti converter. *Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC)*, 2012.

[Slo13]   Guy Slowik. What is magnetic resonance imaging (mri)?, 2013. `http://ehealthmd.com/content/what-magnetic-resonance-imaging-mri`.

[Str06]   Stephen C Strother. Evaluating fmri preprocessing pipelines. *Engineering in Medicine and Biology Magazine, IEEE*, 25(2):27–41, 2006.

[TMT+02]   Jody Tanabe, David Miller, Jason Tregellas, Robert Freedman, and Francois G Meyer. Comparison of detrending methods for optimal fmri preprocessing. *NeuroImage*, 15(4):902–907, 2002.

[VdMH08]   Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(2579-2605):85, 2008.

[VRN+11]   Vincent Q Vu, Pradeep Ravikumar, Thomas Naselaris, Kendrick N Kay, Jack L Gallant, and Bin Yu. Encoding and decoding v1 fmri responses to natural images with sparse nonparametric models. *The annals of applied statistics*, 5(2B):1159, 2011.

Non-exclusive licence to reproduce thesis and make thesis public

I, Martin Loginov (date of birth: 14th of July 1987),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1 reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2 make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Beyond decoding: representational similarity analysis on fMRI data

supervised by Raul Vicente Zafra

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 20.05.2015