Methods and Applications of Multivariate Pattern Analysis in Functional MRI Data Analysis

by

Yash Shailesh Shah

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in the University of Michigan
2015

Doctoral Committee:

Associate Research Scientist Scott J. Peltier, Co-Chair
Professor Douglas C. Noll, Co-Chair
Research Associate Professor Luis Hernandez-Garcia
Assistant Professor Zeeshan Syed
Professor Jon-Kar Zubieta

DEDICATION

To my loving family

ACKNOWLEDGEMENTS

This thesis marks the culmination of a journey which would not have been possible without the contributions of a number of individuals. I would like to express my gratitude to all of them for their numerous sacrifices, encouragement and immense support.

First and foremost, I express my sincere gratitude to Dr. Scott Peltier and Dr. Douglas Noll for being such incredible advisors to me during my graduate studies at the University of Michigan. Scott has been my mentor, advisor, friend and teacher. Most importantly, he has seamlessly transitioned between all these varied roles at the right times. I deeply appreciate his patience, endless support and vast amount of knowledge that he has shared with me. Doug has had a great influence over my life. I have always looked up to him for his composed, humble and warm nature. I cannot recount the number of times he has provided me with perfect guidance and pointed me in the right direction when I was seeking advice. I am glad I knocked on your door years ago without fixing an appointment and thanks for letting me in anyway.

Working with Dr. Luis Hernandez-Garcia has been the most fun and memorable experience of my doctoral studies. Someday, I wish to imitate his immense dedication and enthusiasm towards work. Scott, Doug and Luis are all great teachers and their mentorship has molded me into the person that I am today. I have been very fortunate to gain knowledge under their tutelage.

I would also like to thank the other members of my committee, Dr. Zeeshan Syed and Dr. Jon-Kar Zubieta for their constructive feedback of my work that has helped improve the quality of this dissertation. Thank you for your time and effort.

Special thanks to Dr. Zahi Karam for steadying my rocky boat, strengthening my sails and realigning them to set me on the right course again. Without your technical expertise and words of encouragement my journey would not have been as smooth. Thank you.

I would like to acknowledge with much appreciation the support of Maria Steele for her tremendous help in handling administrative obstacles in times of trouble and Chuck Nicholas for being a patient teacher especially with regards to all computer issues.

Thank you, Dr. Jon-Fredrik Nielsen for always being supportive and giving me helpful advice, especially, on writing and presentation skills. I also thank Krisanne Litinas for being a great teammate and friend and for having a mini script coded for everything that made all our lives easier. Thanks to Keith Newnham, Ryan Smith and Ruth Halsey for making fMRI lab feel like home. I can never underestimate the contribution of all my labmates and friends: Kiran, Valur, Christie, Daehyun, Hesam, Feng, Angela, Vivek, Hao, Matt, Alan, Steven, Kathleen, Sydney, Yang and Gopal to name a few. With them I have found inspirational conversations, random discussions and close friendships.

Outside the fMRI lab, I have been extremely fortunate to have great friends who provided me with the right number of distractions and then brought me back on track too. With them I have shared happiness and tears. I cannot possibly name all my friends who have contributed and motivated me along my journey, but I especially thank Rahul, Sneha, Amritha, Pinaki, Uttara, Ravtaj, Shubhi, Maya, Sharath, Ashray, Vivek, Kartik and Neha. They have truly made me feel at home, away from home.

Most of all, I want to thank my parents, Ashita and Shailesh. Thank you, mom, for your immense unconditional love and belief in me. Dad, thanks for your words of wisdom and undying support. My brother, Jay, has always held the torch to direct me in the darkness and my

sister-in-law, Neha, has never failed to cheer me up when I needed it. I can never express in words, the contribution my brother has had in shaping my life. Thank you, Bhai.

I thank each one of you for being co-passengers in my wonderful journey and forging unforgettable memories together.

Yash Shah

Ann Arbor, Michigan

January 2015

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF ABBREVIATIONS

fMRI	Functional Magnetic Resonance Imaging
MRI	Magnetic Resonance
SVM	Support Vector Machine
SVR	Support Vector Regression
PCASL	Pseudo-Continuous Arterial Spin Labeling
AVAST	Arterial Volume-weighted Arterial Spin Tagging
RSFC	Resting State Functional Connectivity
RF	Radio Frequency
MR	Magnetic Resonance
BOLD	Blood Oxygenation Level Dependent
HR	Hemodynamic Response
CBF	Cerebral Blood Flow
CBV	Cerebral Blood Volume
ASL	Arterial Spin Labeling
SNR	Signal-to-Noise Ratio
GLM	General Linear Modeling
MVPA	Multivariate Pattern Analysis
SOM	Self-Organizing Maps
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis

LR	Logistic Regression
ACC	Anterior Cingulate Cortex
aCBV	arterial Cerebral Blood Volume
TE	Time of Echo
TR	Time of Repetition
SPM	Statistical Parametric Mapping
FWHM	Full Width at Half Maximum
TMD	Temporomandibular Disorders
ASD	Autism Spectrum Disorders
DMN	Default Mode Network
PCC	Posterior Cingulate Cortex
AD	Asperger's Disorder
HC	Healthy Control
WASI	Wechsler Abbreviated Scale of Intelligence
DSM	Diagnostic and Statistical Manual
LOOCV	Leave One Out Cross-Validation
MS	Multi-subject
AAL	Automated Anatomical Labeling
OFC	Orbitofrontal Cortex
mOFC	medial Orbitofrontal Cortex
ROI	Regions of Interest
SMA	Supplementary Motor Area
UDP	User Datagram Protocol

ABSTRACT

In spite of the tremendous advances in science and technology, the human brain and its functions are still not completely understood. Functional magnetic resonance imaging (fMRI) is an imaging modality that allows for non-invasive study of brain function and physiology. Thus, fMRI has found many applications in various fields involved in the study of cognition, psychology, psychiatry, neuroscience, etc. Machine learning techniques have gained tremendous interest in recent times for fMRI data analysis. These methods involve learning from numerous examples and then making predictions for new unseen examples. This work addresses the use of machine learning techniques to find and study multivariate patterns in the fMRI brain data.

The two main applications explored in this work include temporal brain-state prediction and subject categorization. The within-subject brain-state prediction setup has been used to compare and contrast three different acquisition techniques in a motor-visual activation study. It has also been implemented to highlight the differences in pain regulation networks in healthy controls and subjects with temporomandibular disorders. Lastly, regression has been used to predict graded fMRI activation on a continuous scale in a motor activation and craving study. The between-subject categorization setup has been used to distinguish between patients with Asperger's disorder and healthy controls.

A major contribution of our work involves a novel multi-subject machine learning framework. This technique helps to learn a model which is based on information acquired from multiple other subjects' data in addition to the subject's own data. This has been used to classify the craving and non-craving brain states of nicotine-dependent subjects, allowing examination of both population-wide as well as subject-specific neural correlates of nicotine craving. A real-time neurofeedback setup was implemented to provide feedback to a subject using their own brain activation data. Subjects can then be trained to self-regulate their own brain activation.

CHAPTER 1

INTRODUCTION

1.1 Overview

The objective of this dissertation is to describe and explore novel methods and applications of multivariate pattern recognition in functional magnetic resonance imaging (fMRI) data analysis. We employ machine learning techniques on the complex, multivariate neuroimaging data to achieve this goal.

Machine learning techniques are a diverse group of methods that enable computers to learn from examples. After examining a set of examples, these algorithms help computers build models that can later be used to classify new unseen data into distinct classes.

Magnetic resonance imaging (MRI) is an imaging modality which uses a superconducting magnet to acquire structural images of the human body. In functional MRI, the MRI equipment can be used to acquire images of the human brain to understand brain function or to localize these functions to certain parts of the brain. Thus, fMRI can help in the study of neural or cognitive processes or to enable brain mapping.

Recently, there has been growing interest in the use of machine learning techniques for analyzing fMRI data. The goal of this endeavor is to extract meaningful information from neuroimaging data. This avenue of inquiry involves finding a pattern, localizing it and then characterizing it.

In this dissertation we have mainly explored multivariate pattern classification for two end goals: 1) classification of a subject's mental states (within-subject classification) or 2) categorization of subjects into healthy controls or patients with disorders (between-subject classification). We have also presented a novel method called multi-subject machine learning for building an informed classifier for an individual which benefits from the information provided by multiple subjects in addition to the subject's own data.

1.2 Thesis organization

The chapters after this introduction are organized as follows:

In chapter 2, we have introduced relevant background and motivation for the rest of the thesis. It poses research questions and provides context for the topics that are covered in the following chapters. We start with introducing core concepts of functional brain imaging and machine learning for multivariate pattern analysis. Since our major application involves the study of neural processes related to cigarette craving, we introduce the problem statement and discuss the motivation behind studying it.

In chapter 3, we discuss the experiment for within-subject temporal brain state classification. Support vector machine (SVM) classification is used to classify brain volumes into task activation or rest states. We compare three different acquisition techniques and highlight the pros and cons of each. A recently developed variant of pseudo-continuous arterial spin labeling (PCASL) named arterial volume-weighted arterial spin tagging (AVAST) is shown to perform good classification while exhibiting various desired characteristics. We then show results of a brain state classification experiment of pain vs. rest on TMD subjects and healthy controls. Lastly, we report results of predicting graded fMRI activation using support vector regression (SVR).

In chapter 4, we discuss a between-subject classification scheme. Resting state functional connectivity (RSFC) maps are used for this purpose. SVM classification is used to predict the disease state of RSFC maps as belonging to healthy control subjects or subjects diagnosed with Asperger's syndrome.

In chapter 5, we present methods that are used to image craving and non-craving brain states in nicotine dependent subjects. The novel multi-subject machine learning approach used to build the classifier is shown to perform better than traditional approaches and also reconfirms the subjective nature of nicotine craving. We then discuss the real-time neurofeedback setup developed and used in our study.

Finally, chapter 6 presents a summary of knowledge discovery and contributions of this thesis and proposed future work.

CHAPTER 2

BACKGROUND AND MOTIVATION

2.1 Magnetic resonance imaging

Magnetic resonance imaging (MRI) is a technique based on the principles of nuclear magnetic resonance used extensively as a medical imaging tool to acquire structural images of the human body. Typically, it involves the use of a large superconducting magnet, radio frequency (RF) transmit and receive coils and gradient coils to manipulate atomic spins in the human body. Due to their presence in water molecules, hydrogen atoms are the most common atoms in the human body. In normal circumstances, these hydrogen nuclei are randomly oriented but in the presence of the superconducting magnet, they align themselves in the direction of the strong magnetic field. The RF transmit coil applies a radio frequency magnetic pulse at the correct frequency to disturb the hydrogen nuclei from their equilibrium and the RF receive coil detects the faint MR signal emitted by the nuclei during their return to the equilibrium state. The gradient coils are used to produce variations in the main magnetic field to enable localization of the signal. The final MR image is a map of the distribution of MR signal obtained. By careful design of RF and gradient pulses and delays, subtle changes in anatomy can be detected using this method. Thus, MRI can be used for structural mapping. [1, 2, 3, 4].

Among many other features, few of the most vital characteristics that have propelled MRI as a preferred method over other medical imaging modalities include its non-invasive nature and lack of ionizing radiation.

2.2 Functional magnetic resonance imaging

Functional magnetic resonance imaging (fMRI) is a type of specialized MRI scan of the brain or spinal cord. It involves the use of MRI equipment to capture a sequential series of MR images. This time series of images can then be used to detect regional changes in cerebral metabolism, blood flow or blood volume in response to task activation or during rest, thus enabling functional mapping of the brain [5, 6, 7]

Functional MRI is becoming a preferred method for studying the functioning of a normal, diseased or injured brain. It is also used for finding distinguishing characteristics in the brain that enable the identification of disorders, as well as for assessing the potential risks of surgery or other invasive treatments of the brain. A noteworthy feature of fMRI is that it does not involve the use of an external contrast agent but most commonly depends on contrast provided by blood oxygenation, perfusion, diffusion, etc.

2.2.1 BOLD fMRI

The most popular technique in fMRI utilizes blood oxygenation level dependent (BOLD) contrast which was first reported in [8]. The human brain requires a steady supply of oxygen for its proper functioning which is provided by the hemoglobin molecules in the blood. Hemoglobin is the iron-containing oxygen-transporting protein present in red blood cells. It contains iron atoms to which oxygen atoms may or may not be bound. When oxygen atoms are bound to hemoglobin (oxyhemoglobin), then the iron is shielded by oxygen and thus it exhibits diamagnetic properties, whereas, deoxyhemoglobin is paramagnetic. Oxygenated blood is diamagnetic whereas deoxygenated blood is paramagnetic, and this difference in their magnetic susceptibility is exploited to generate detectable changes in susceptibility-weighted MR images using the BOLD fMRI technique [9].

When a neuron is activated, its metabolic demand consumes oxygen in the nearby region and thus there is a momentary decrease in blood oxygenation. This increase in deoxyhemoglobin as compared to oxyhemoglobin causes the initial dip in fMRI signal. In case of a persistent demand for oxygen, the flow of oxygenated blood to the area increases and overcompensates for the initial dip. Since the supply of oxyhemoglobin is greater than the demand, this results in an increase in the fMRI signal. After the demand for blood ceases, a post-stimulus undershoot is sometimes observed in the fMRI signal before it finally returns to baseline. This phenomenon of change in blood oxygenation in response to neural activity is called the hemodynamic response (HR) and it characterizes the shape of the fMRI signal in an activated brain region. Thus by measuring the BOLD signal change, fMRI technique can identify a brain region that is activated.

2.2.2 ASL fMRI

Because of the widespread use of BOLD contrast to acquire fMRI brain images, fMRI has become synonymous with BOLD imaging. Other methods that are used to map brain function typically involve measurement of cerebral blood flow (CBF) or cerebral blood volume (CBV). Arterial spin labeling (ASL) is one such method that utilizes NMR-labeled water as an intrinsic tracer to measure CBF as an indicator of fMRI activity.

In ASL techniques, CBF is estimated by measuring the concentration of the tracer as it passes through a tissue of interest. Instead of injecting an external tracer, RF electromagnetic pulses are used to invert the magnetization of hydrogen protons in the blood to magnetically label water and create an endogenous tracer. After allowing a short period of time for tracer to reach the tissue of interest, an image is collected with the tracer (tag image). This process is repeated without applying an RF pulse and an image is collected without tracer (control image). CBF can then be quantified by measuring the signal change due to the tracer by subtracting the

tag image from a control image. Due to the short relaxation time of arterial blood, the tracer decays rapidly and the signal-to-noise ratio (SNR) is very low. Since two images need to be acquired and then subtracted to obtain one signal, ASL techniques also suffer from low temporal resolution.

The signal acquired from BOLD imaging cannot be easily quantified and it also lacks a clear physiological interpretation. The BOLD signal is sensitive to field inhomogeneities caused by the differences of magnetic susceptibility of air and tissue which may result in local image distortions and signal losses [10, 11]. The within-session slow scanner drift which has been observed in most studies using BOLD imaging [12], makes it impractical for studies involving neural processes with long activation periods.

Arterial spin labeling holds answers for many of these shortcomings of the BOLD imaging technique. Since it involves the measurement of CBF, it can be directly quantified in terms of flow units. Thus, ASL techniques can be used for longitudinal as well as multisite studies. ASL is not as severely affected by field inhomogeneities and thus can be effectively used in brain regions close to air-tissue interfaces as well. Since ASL is a subtraction technique, the effects of low frequency scanner drift are diminished considerably. This advantage makes ASL techniques an ideal choice for imaging studies of neural processes with long activation periods.

2.3 Multivariate Pattern Analysis using Machine Learning in fMRI

Most fMRI studies were traditionally analyzed using a univariate approach called Statistical Parametric Mapping [13], which looks at the extent to which a voxel's time series is explained by a few regressors in a general linear model. Several other whole-brain mass univariate analysis techniques have been implemented in the past to detect patterns in the brain imaging data [14, 15]. However, univariate approaches treat each voxel's time course separately and do not take into account inter-regional correlations which may be of importance in studies that are directed at exploring various neural systems associated with a particular aspect of brain function. Also, they do not offer the prospect of employing a predictive learning approach which may be of significant diagnostic relevance.

MVPA applies sophisticated machine learning techniques to the complex patterns generated by the very large number of features i.e. voxel intensities [16]. Machine learning and pattern recognition techniques are being increasingly employed in fMRI data analysis [17]. By taking into account the full spatial pattern of brain activity measured simultaneously at many locations, these methods allow detecting subtle localized effects that may remain unidentified with conventional univariate statistical methods.

Machine learning techniques are used to find patterns in the data. As suggested by the name, machine learning methods are algorithms that allow a computer to learn. In typical fMRI applications, the machine learning algorithms are used to learn a relationship between brain volumes and labels. This learned functional relationship is used to form a model which is then used to predict the unseen labels for a new test data set. Thus they facilitate a classifier-based predictive learning framework. One sample of the input data is called an example or an instance, and every instance is described by a set of features.

A lot of different machine learning techniques have been used for multivariate pattern analysis in numerous functional MRI studies to investigate different neural processes. Support vector machines and linear discriminant analysis were applied to successfully classify patterns of fMRI activation observed due to the visual presentation of pictorial cues of various categories of objects in [18]. It was demonstrated that fMRI activity patterns in early visual areas contain detailed orientation information that can reliably predict subjective perception using linear SVM in [19]. In another study [20], support vector machines were used for temporal classification of data in a blocked design experiment and reported results with the use of linear as well as polynomial kernels.

In [21], the authors have successfully applied machine learning techniques to implement lie-detection using non-linear SVM. They compare this performance with Fisher's linear discriminant analysis using the same data and report poorer classification accuracy. In [22], the authors train a model of observed fMRI data associated with viewing several dozen concrete nouns from a large text corpus. Once trained, the model predicts fMRI activation for thousands of other concrete nouns with highly significant accuracies.

A variety of unsupervised learning methods have also been used for exploratory analysis of fMRI data. In [23], an approach for optimization of PCA is reported, whereas, the use of clustering methods for fMRI data is discussed in [24]. K-means clustering has also been used with resting state fMRI time series data in [25] whereas, ICA has been used in multiple studies mainly to investigate differences in fMRI activations of healthy controls and disordered populations [26, 27, 28].

In fMRI data analysis, machine learning can be used for one of two major applications:

(1) Classification of a subject's mental states (within-subject classification)

When posed as a temporal brain-state classifier, they learn a functional relationship between brain response patterns and a perceptual, cognitive or behavioral state of a subject expressed in terms of a label, which may be assigned discrete or continuous values to be used in a classification or regression model, respectively. Each brain volume acquired during a separate time point acts as an instance and the intensities of all voxels in that volume act as features.

(2) Categorization of subjects into healthy controls or patients with disorders (between-subject classification)

When used for subject categorization, the algorithm learns a functional relationship between a subject's data and a class label which describes the disease state of that subject (for example: +1=healthy control, -1=patient with disorder). For each subject, the temporal dimension can be reduced by using correlation and creating functional connectivity maps. Each subject's functional connectivity map can then act as an instance and the correlation values act as features.

The following few figures show an illustration of classifier-based analysis. Figure 2.1 shows an example of brain volume acquired and displays one axial slice. All the voxels in all such slices are concatenated into a single row to form one example. Each voxel represents a separate feature or attribute of the example. As shown in figure 2.2, each example is assigned a label. The entire dataset comprises of a collection of many such labeled examples.

In within-subject classification, as illustrated in figure 2.3, the different examples are brain volumes acquired at different time points and the labels are assigned depending on subject's brain state as dictated by the experimental design. As shown in figure 2.4, in the case of between-subject classification, the time dimension is reduced by computing the correlation

between time course of each voxel and that of the seed time course and the different examples are correlation maps of different subjects. They are assigned labels depending on the subjects' disease state.

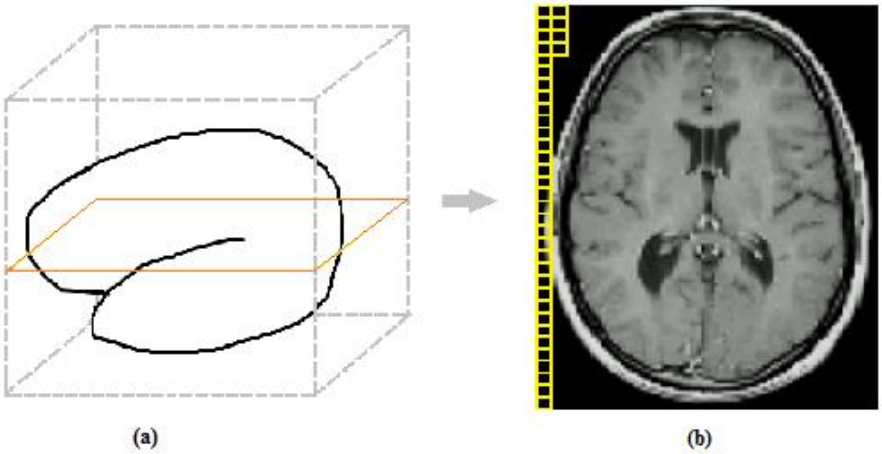

Figure 2.1: Illustration showing brain volume and voxels.
(a) Brain volume acquired showing one axial slice,
(b) One slice showing voxels that act as features.

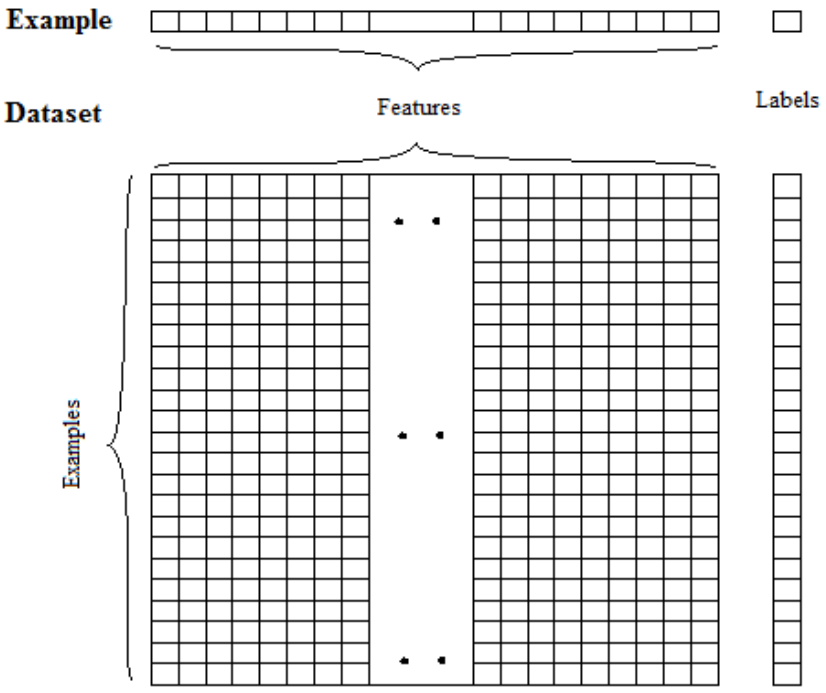

Figure 2.2: Illustration of classifier based analysis.
Each row represents a single example with a label assigned to it.
A dataset comprises of a collection of such examples.

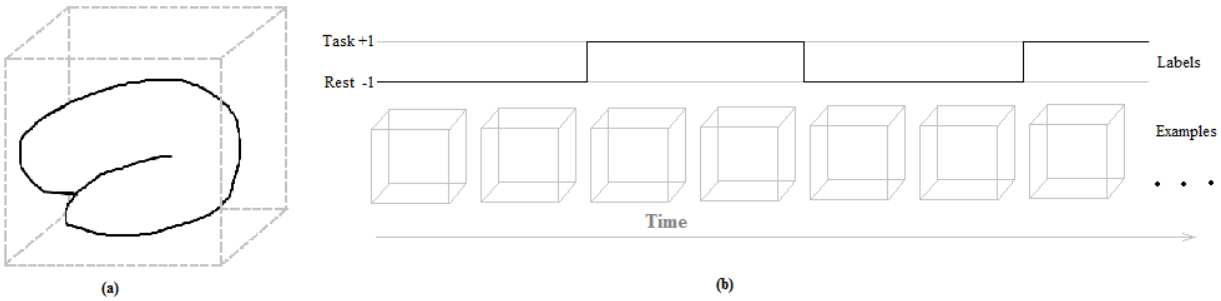

Figure 2.3: Schematic of within-subject classifier.
(a) Brain volume acquired at each time point.
(b) Many such volumes are acquired across time and are assigned labels depending on brain state (Task = +1, Rest = -1)

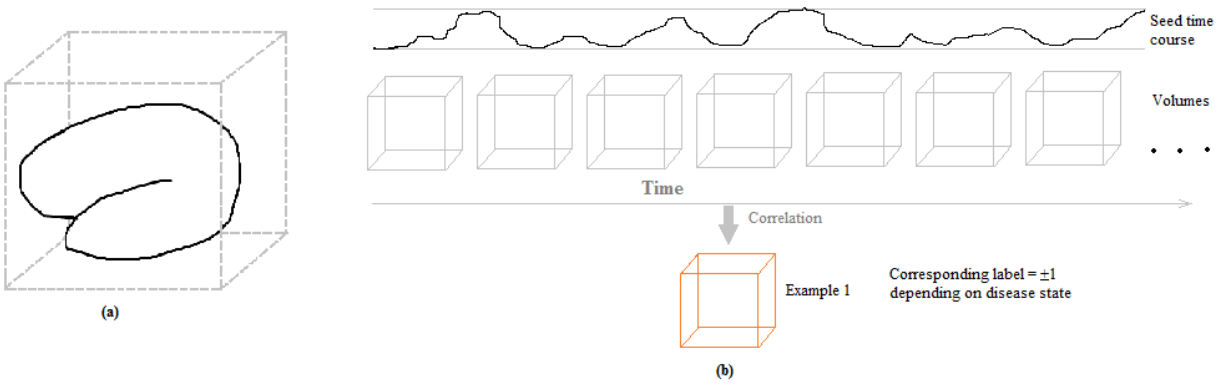

Figure 2.4: Schematic of between-subject classifier.
(a) Brain volume acquired at each time point.
(b) Correlation is computed between time course of each voxel and time course of seed voxel.
One correlation map is obtained per subject and assigned label depending on their disease state (Healthy control = +1, Patient with disorder = -1)

Machine learning techniques can be classified into different groups based on many different criteria. When this categorization is done based on learning style, then these techniques can be broadly classified into two main divisions. (1) Unsupervised and (2) Supervised.

2.3.1 Unsupervised learning

Unsupervised machine learning techniques are data-driven and do not involve any explicit mention of class labels during training. They form models by identification of natural groups or patterns within data. Unsupervised machine learning can be used to group the input data into classes on the basis of their statistical properties alone. Clustering which involves grouping data into classes based on some measure of inherent similarity is an example of unsupervised learning. Some examples of unsupervised methods used for fMRI data analysis include k-nearest neighbor, self-organizing maps (SOM), k-means clustering [25], independent component analysis (ICA) [26, 27, 28].

2.3.2 Supervised learning

Supervised learning is based on determining a mapping between particular features of the data to given target labels. It is divided into two phases, training phase and testing phase. During the training phase, a set of data points belonging to the training data set is used to estimate the parameters of a model relating the features to the target labels. Once the parameters are learned, the model can then be applied to predict the label of a previously unseen data point belonging to the test data set. If the target labels comprise a set of discrete classes, the supervised learning problem is referred to as classification and as regression when the target labels assume continuous values. Different supervised algorithms make a different set of assumptions about the data and have a different means of learning a model. Some of the supervised machine learning

techniques used for neuroimaging data analysis include support vector machine (SVM) [20, 29], linear discriminant analysis (LDA) [29] and logistic regression (LR) [30].

2.3.2.1 Support vector machine classification

Support vector machines are a powerful set of such machine learning methods that can be used to analyze data and recognize patterns [31, 32, 33]. In this study, support vector machines were chosen as the choice of machine learning technique for multiple reasons. They can be implemented for use in classification as well as regression analysis. Another vital characteristic is that the SVM algorithm seeks a maximum margin separating hyperplane, thus making it resilient to overfitting. This means that they provide better generalization, allowing the best prediction accuracy for previously unseen test data. Also, when a linear kernel is used, they allow the possibility of generating discrimination maps so that they can be visualized in the original data space. Thus, SVMs help not only in effective pattern discrimination but also pattern localization in the given data. SVMs are a preferred method for learning when the data has many more features than the number of examples which is the typical case for fMRI experiments.

When implemented as a classifier, support vector machines learn from a set of training examples, which are labeled as belonging to either of two classes, and build a model. During the testing phase, this model then predicts whether an unseen test example belongs to one class or the other in an extremely fast process. This provides a much needed advantage in order to implement real-time neurofeedback.

Basically, a support vector machine constructs a hyperplane in high-dimensional space which is capable of separating examples which belong to either of two classes. In order to maintain the generalization error to a minimum, a maximum margin classifier is sought, i.e. the

hyperplane which ensures the largest distance to the nearest training point in either class. These nearest training points that rest on the boundary are termed the support vectors.

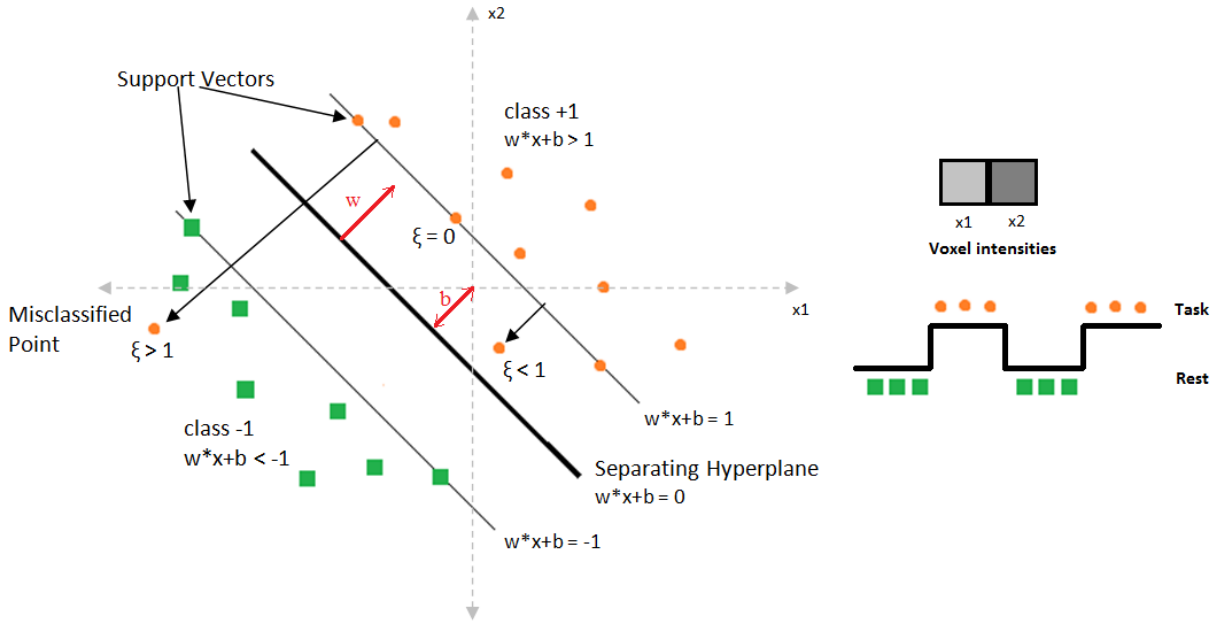

Figure 2.5: Support vector machine classification example.

Figure 2.5 attempts to explain the working of support vector machine classification with the help of a simple example in which a line is sought to classify the samples belonging to either of two classes (orange circles and green squares). These samples are voxel intensities of a hypothetical two-voxel brain at different time points in either task or rest condition. For simplicity, consider a simple example of a two voxel brain with x1 and x2 being the gray scale intensities of these two voxels. The paradigm design is briefly illustrated on the right hand side. Green squares represent rest whereas orange circles represent task condition. Alternating blocks of rest and task conditions are presented to the subject and data from this hypothetical two voxel brain is acquired at each time point. A maximum margin separating hyperplane is sought to classify these brain data collected at every time point into rest (-1) or task (+1) condition. Once

this classifying hyperplane is constructed, new test data can be easily classified depending on which side of the hyperplane they lie.

The separating hyperplane generated by the SVM algorithm is orthogonal to the weight vector **w** which defines the direction in which the examples of the two classes differ most from one another. Thus, the classifier is parameterized by **w,** which can be solved for by using the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{w}} \| \mathbf{w} \|^2 + C \sum_{i=1}^{n} \xi_i \\ &s.t. \; \forall i \in \{1, 2, ..., n\} : \\ &\mathbf{y}_i \mathbf{w}^T \mathbf{x}_i \geq 1 - \xi_i \; \& \; \xi_i \geq 0 \end{aligned} \qquad (1)$$

where **w** is the normal vector to the hyperplane, $\mathbf{y}_i$ are the known input class labels, $\mathbf{x}_i$ are the input feature vectors, C is a tradeoff parameter and ξ are non-negative slack variables.

Some misclassifications, as represented by the two orange dots on the wrong side of the boundary, may also be permitted. But these misclassifications are penalized in the optimization problem depending on the distances represented by the slack variables (ξ). C is the trade-off parameter used to regulate the number of misclassifications to allow. A very high value of C penalizes all misclassifications heavily and thus builds a classifier which might overfit the data. On the other hand, a low value of C might permit too many misclassifications and an inept model might be obtained. Thus, cross validation may be performed to choose the best value of C.

2.3.2.2 Support vector regression

When implemented for regression analysis, support vector machines compute functional approximation. Given a set of input training examples with real-valued labels, it tries to fit a function to these training samples and then when presented with an unseen test example, it

calculates the corresponding output of the function. In regression analysis, Vapnik's ε-insensitive loss function is used to define an ε-tube [31]. As shown in figure 2.6, if the measured value is within the defined ε tube, the loss associated with it is zero and for all other predicted points outside the tube, this loss is equal to difference between the value and the radius ε of the tube.

Figure 2.7 illustrates an example of support vector regression. The epsilon tube is seen as a margin in two dimensions. ξ and ξ* are the losses associated with those particular samples as defined by Vapnik's loss function.

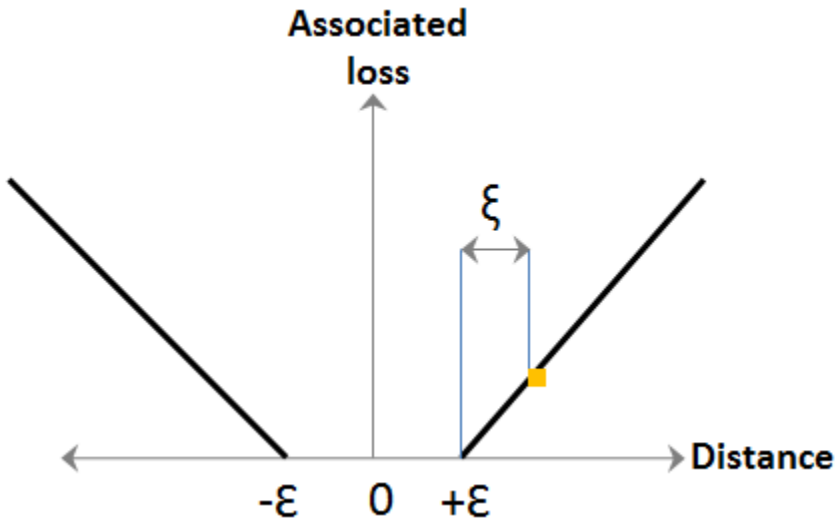

Figure 2.6: Vapnik's ε-insensitive loss function.

The corresponding optimization problem for SVR is given by:

$$\begin{aligned} &\min_{\mathbf{w}} \| \mathbf{w} \|^2 + C \sum_{i=1}^{n} \xi_i \\ &s.t. \ \forall \ i \in \{1, 2, \ldots, n\}: \\ &\mathbf{y}_i - \mathbf{w}^T \mathbf{x}_i - b \leq \varepsilon + \xi_i \ \ \& \\ &\mathbf{w}^T \mathbf{x}_i + b - \mathbf{y}_i \leq \varepsilon + \xi_i * \end{aligned} \qquad (2)$$

where **w** is the normal vector to the hyperplane, $\mathbf{y}_i$ are the known input class labels, $\mathbf{x}_i$ are the input feature vectors, C is a tradeoff parameter, ξ and $\xi*$ are non-negative slack variables and ε defines the tube.

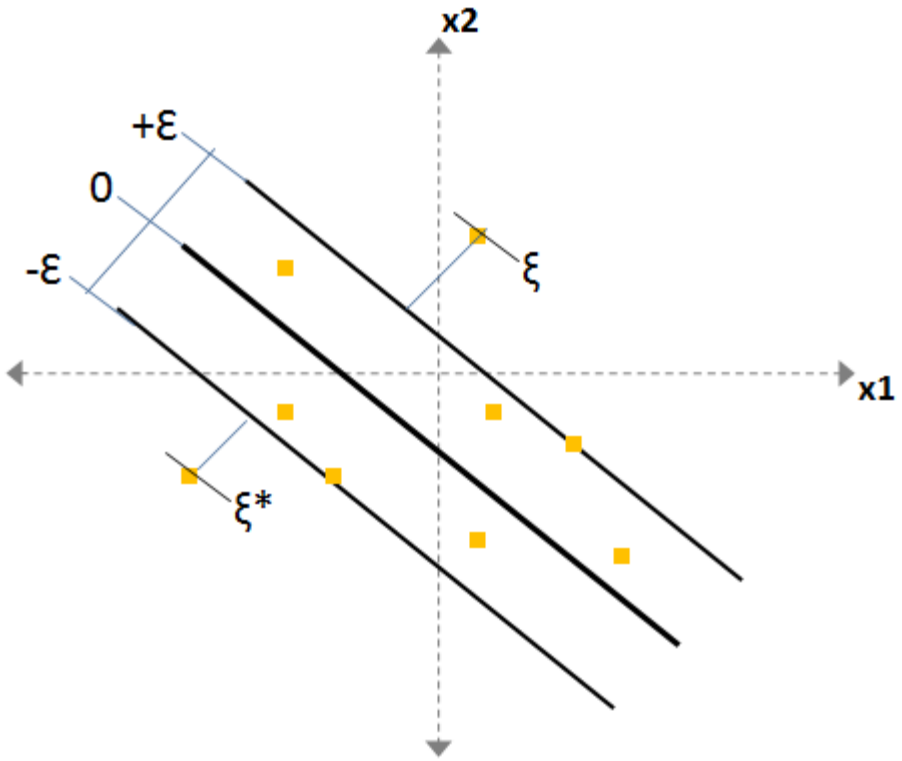

Figure 2.7: Support vector regression example.

Another important point to note is that kernel mapping techniques can be employed with support vector machines so as to construct a linear classifier to classify the input vectors in a higher-dimensional feature space instead of the input space, without having to explicitly convert them into this feature space [32]. Figure 2.8 shows a schematic diagram explaining this kernel mapping with support vector machines. Kernel mapping may be employed in order to perform classification as well as regression of the data. But in most of our examples, the number of features (voxels) and hence the dimensionality of data is much higher than the number of instances and hence the linear kernel is sufficient for finding a separating hyperplane [33].

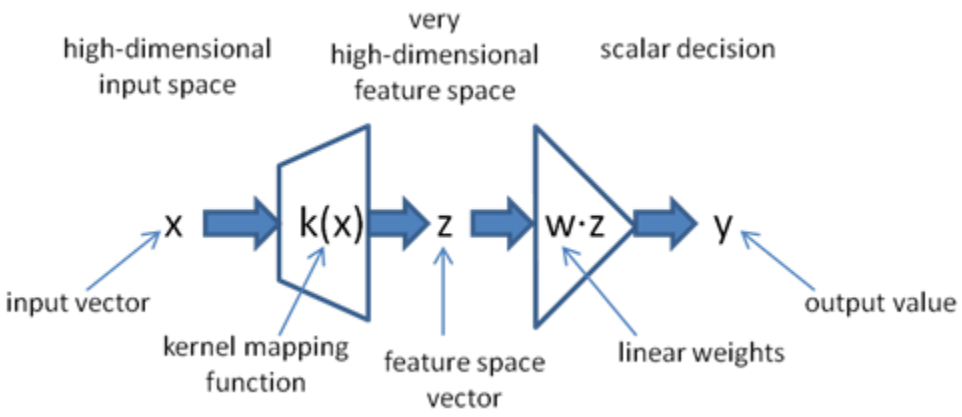

Figure 2.8: Kernel mapping technique schematic.

2.4 Smoking

2.4.1 Addiction and effects

Self-regulation of nicotine craving was chosen ahead of any other addictive drug due to a number of reasons. Smoking is a major addictive public health problem [34] and the number one cause of preventable morbidity and mortality in the U.S. [35]. Also, due to the ease of availability of cigarettes, there is a high prevalence of nicotine dependence and thus smokers are much easier to recruit as compared to illegal drug users. Recent neuroimaging studies have examined regional activity responses in correlation with craving for cigarettes and reported an increase in brain activation in the prefrontal cortex [36, 37, 38], thalamus [36, 37, 39, 40] and visual system [37, 39]. Another inhalation study reported an increase in frontal lobe and thalamic blood flow in smokers who smoked a cigarette [41]. These findings have reinforced the idea that nicotine has direct effects on neural function [42, 43]. Nicotine dependence is a prototypical form of drug dependence and findings from this project may encourage similar research with other forms of drug dependence.

2.4.2 Nicotine craving

Craving can be a persistent and disturbing feature of addiction that can contribute to drug-seeking behavior or relapse in subjects trying to remain abstinent. It describes an irresistible need for drug intake that eventually leads to dependent behavior. Craving is thought to originate from positive-reinforcing properties of the drug [44] or withdrawal-related negative-reinforcing properties [45]. Frequent smokers as well as consumers of alcohol and illegal drugs report moderate craving for the drugs they use [46].

Cocaine users report experiencing a strong desire for the drug when exposed to cues associated with it. Nicotine dependence level is associated with greater BOLD fMRI activation [47, 48] and cigarette craving [49] in response to smoking cues. Mental imagery of smoking related material (e.g. audio scripts or visual cues) produces a consistent profile of craving and mood effects compared to material devoid of smoking-related cues. This suggests the efficacy of using auditory and visual paradigms to elicit craving and the prospect of identifying crave or no-crave state of a brain based on their functional MRI data when presented with such cues.

2.4.3 Neurofeedback and self-regulation

Biofeedback is a technique by which subjects' bodily processes may be measured and then presented back to them immediately to enable increased awareness and deliberate control of the related cognitive or physiological activity. Multiple groups have investigated and practiced this technique in the past. Recent advances in acquisition techniques, computational power and algorithms have increased the sensitivity and speed of fMRI significantly, making real-time acquisition, analysis and display of fMRI data feasible. These developments together facilitate the possibility of a real-time fMRI neurofeedback system which enables online feedback of the BOLD signal. This system can enable subjects to learn the self-regulation of local brain activity with a high spatial and temporal resolution and whole-brain coverage. Recent work using real-time fMRI establishes that an individual can learn to directly control activation of localized brain regions that are associated with pain perception and regulation [50]. In another study, signal from the amygdala was used to enhance the effect of a sadness induction paradigm [51]. Other brain regions have also been shown to respond to real-time fMRI feedback including the auditory [52], and motor cortex in a finger tapping paradigm [53], rostral ACC [54, 55] and anterior insular cortex [56]. This technique may be extended to control the activation of brain regions that may

be involved in craving. This approach might provide control over the neurophysiological mechanisms of craving in nicotine dependent subjects and may thus prove to be a particularly effective neurofeedback method.

2.5 Overall Significance

There is an ever-increasing rise in the number of mortalities due to the negative health effects of cigarette smoking. Because of the addictive nature of nicotine present in cigarettes, smokers find it extremely difficult to quit and there is a large number of cases of relapse due to craving. Hence, it would be of tremendous significance to devise a technique that could assist nicotine dependent subjects to control their craving or their urge to smoke. As described in this dissertation, real-time fMRI neurofeedback proposes one such possible technique. This thesis explores the possibility of exploiting the fast nature of powerful machine learning techniques such as SVM for this application.

Also, in this dissertation, the possibility of using the machine learning techniques for learning brain state predictors and models that can delineate between healthy and diseased populations has also been investigated.

REFERENCES:

[1] D. G. Nishimura, *Principles of magnetic resonance imaging*. Unpublished textbook, 1996.

[2] R. Damadian, "Tumor Detection by Nuclear Magnetic Resonance," *Science1*, vol. 171, no. 3976, pp. 1151–1153, 1971.

[3] P. C. Lauterbur, "Image formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonace," *Nature*, vol. 242, pp. 190–191, 1973.

[4] P. Mansfield, "Multi-planar image formation using NMR spin echoes," *J. Phys. C. Solid State Phys.*, vol. 10, pp. L55–L58, 1977.

[5] R. B. Buxton, *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge University Press, 2002.

[6] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*. 2009.

[7] M. E. Raichle and M. a Mintun, "Brain work and brain imaging.," *Annu. Rev. Neurosci.*, vol. 29, pp. 449–76, Jan. 2006.

[8] S. Ogawa, T. M. Lee, A. S. Nayak, and P. Glynn, "Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields," *Magn. Reson. Med.*, vol. 14, no. 1, pp. 68–78, Apr. 1990.

[9] N. K. Logothetis, "The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal.," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 357, no. 1424, pp. 1003–37, Aug. 2002.

[10] N. Weiskopf, C. Hutton, O. Josephs, and R. Deichmann, "Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T.," *Neuroimage*, vol. 33, no. 2, pp. 493–504, Nov. 2006.

[11] B. P. Sutton, S. Member, D. C. Noll, J. A. Fessler, and S. Member, "Fast , Iterative Image Reconstruction for MRI in the Presence of Field Inhomogeneities," vol. 22, no. 2, pp. 178–188, 2003.

[12] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. Yang, J. H. Duyn, and J. A. Frank, "Investigation of Low Frequency Drift in fMRI Signal," vol. 533, pp. 526–533, 1999.

[13] K. Friston, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. 2006.

[14] S. M. Smith, “Overview of fMRI analysis,” *Br. J. Radiol.*, vol. 77, no. suppl_2, pp. S167–S175, Dec. 2004.

[15] W.L. Luo and T. E. Nichols, “Diagnosis and exploration of massively univariate neuroimaging models,” *Neuroimage*, vol. 19, no. 3, pp. 1014–1032, Jul. 2003.

[16] E. Formisano, F. De Martino, and G. Valente, “Multivariate analysis of fMRI time series: classification and regression of brain responses using machine learning.,” *Magn. Reson. Imaging*, vol. 26, no. 7, pp. 921–34, Sep. 2008.

[17] F. Pereira, T. Mitchell, and M. Botvinick, “Machine learning classifiers and fMRI: a tutorial overview.,” *Neuroimage*, vol. 45, no. 1 Suppl, pp. S199–209, Mar. 2009.

[18] D. D. Cox and R. L. Savoy, “Functional magnetic resonance imaging (fMRI) ‘brain reading’: detecting and classifying distributed patterns of fMRI activity in human visual cortex,” *Neuroimage*, vol. 19, no. 2, pp. 261–270, Jun. 2003.

[19] Y. Kamitani and F. Tong, “Decoding the visual and subjective contents of the human brain.,” *Nat. Neurosci.*, vol. 8, no. 5, pp. 679–85, May 2005.

[20] S. M. LaConte, S. Strother, V. Cherkassky, J. Anderson, and X. Hu, “Support vector machines for temporal classification of block design fMRI data.,” *Neuroimage*, vol. 26, no. 2, pp. 317–29, Jun. 2005.

[21] C. Davatzikos, K. Ruparel, Y. Fan, D. G. Shen, M. Acharyya, J. W. Loughead, R. C. Gur, and D. D. Langleben, “Classifying spatial patterns of brain activity with machine learning methods: application to lie detection.,” *Neuroimage*, vol. 28, no. 3, pp. 663–8, Nov. 2005.

[22] T. M. Mitchell, S. V Shinkareva, A. Carlson, K.M. Chang, V. L. Malave, R. A Mason, and M. A. Just, “Predicting human brain activity associated with the meanings of nouns.,” *Science*, vol. 320, no. 5880, pp. 1191–5, May 2008.

[23] L. K. Hansen, J. Larsen, F. Å. Nielsen, S. C. Strother, E. Rostrup, R. Savoy, N. Lange, J. Sidtis, C. Svarer, and O. B. Paulson, “Generalizable Patterns in Neuroimaging : How Many Principal Components ?,” vol. 544, pp. 534–544, 1999.

[24] C. Goutte, P. Toft, E. Rostrup, F. A. Nielsen, and L. K. Hansen, “On Clustering fMRI Time Series,” *Neuroimage*, vol. 9, no. 3, pp. 298-310, 1999.

[25] A. Mezer, Y. Yovel, O. Pasternak, T. Gorfine, and Y. Assaf, “Cluster analysis of resting-state fMRI time series.,” *Neuroimage*, vol. 45, no. 4, pp. 1117–25, May 2009.

[26] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, “Investigations into resting-state connectivity using independent component analysis.,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 360, no. 1457, pp. 1001–13, May 2005.

[27] V. D. Calhoun, P. K. Maciejewski, G. D. Pearlson, and K. A Kiehl, "Temporal lobe and 'default' hemodynamic brain modes discriminate between schizophrenia and bipolar disorder.," *Hum. Brain Mapp.*, vol. 29, no. 11, pp. 1265–75, Nov. 2008.

[28] S. A. R. B. Rombouts, J. S. Damoiseaux, R. Goekoop, F. Barkhof, P. Scheltens, S. M. Smith, and C. F. Beckmann, "Model-free group analysis shows altered BOLD FMRI networks in dementia.," *Hum. Brain Mapp.*, vol. 30, no. 1, pp. 256–66, Jan. 2009.

[29] J. Mourão-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, "Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data.," *Neuroimage*, vol. 28, no. 4, pp. 980–95, Dec. 2005.

[30] F. Pereira and M. Botvinick, "Information mapping with pattern classifiers: a comparative study.," *Neuroimage*, vol. 56, no. 2, pp. 476–96, May 2011.

[31] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag NY, 2000.

[32] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[33] L. Hamel, *Knowledge Discovery with Support Vector Machines*. 2009.

[34] P. Jha, M. K. Ranson, S. N. Nguyen, and D. Yach, "Estimates of Global and Regional Smoking Prevalence in 1995, by Age and Sex," *Am. J. Public Health*, vol. 92, no. 6, pp. 1002–1006, Jun. 2002.

[35] A. H. Mokdad, J. S. Marks, D. F. Stroup, and J. L. Gerberding, "Actual Causes of Death in the United States , 2000," *JAMA*, vol. 291, no. 10, 2014.

[36] E. A. Stein, J. Pankiewicz, H. H. Harsch, J. K. Cho, S. A. Fuller, R. G. Hoffmann, M. Hawkins, S. M. Rao, P. A. Bandettini, and A. S. Bloom, "Nicotine-Induced Limbic Cortical Activation in the Human Brain : A Functional MRI Study," *Am J Psychiatry*, vol. 155, no. 8, pp. 1009–1015, August 1998.

[37] E. F. Domino, S. Minoshima, S. Guthrie, L. Ohl, L. Ni, R. A. Koeppe, and J. K. Zubieta, "Nicotine Effects on Regional Cerebral Blood Flow in Awake , Resting Tobacco Smokers," *Synapse*, vol. 38, pp. 313–321, 2000.

[38] J. E. Rose, F. M. Behm, E. C. Westman, J. E. Bates, and A. Salley, "Pharmacologic and sensorimotor components of satiation in cigarette smoking," *Pharmacol. Biochem. Behav.*, vol. 76, no. 2, pp. 243–250, Sep. 2003.

[39] E. D. London, M. Ernst, and S. Grant, "Orbitofrontal Cortex and Human Drug Abuse : Functional Imaging," *Cerebral Cortex*, vol. 10, pp. 334–342, March 2000.

[40] J. K. Zubieta, U. Lombardi, S. Minoshima, S. Guthrie, L. Ni, L. E. Ohl, R. A. Koeppe, and E. F. Domino, "Regional cerebral blood flow effects of nicotine in overnight abstinent smokers," *Biol. Psychiatry*, vol. 49, no. 11, pp. 906–913, Jun. 2001.

[41] H. Nakamura, A. Tanaka, Y. Nomoto, Y. Ueno, and Y. Nakayama, "Activation of fronto-limbic system in the human brain by cigarette smoking: evaluated by a CBF measurement," *Keio J Med*, vol. 49, no. 1, pp. A122–4, 2000.

[42] E. F. Domino, L. Ni, Y. Xu, R. A. Koeppe, S. Guthrie, and J. K. Zubieta, "Regional cerebral blood flow and plasma nicotine after smoking tobacco cigarettes.," *Prog. Neuropsychopharmacol. Biol. Psychiatry*, vol. 28, no. 2, pp. 319–27, Mar. 2004.

[43] J. K. Zubieta, M. M. Heitzeg, Y. Xu, R. A. Koeppe, L. Ni, S. Guthrie, and E. F. Domino, "Regional Cerebral Blood Flow Responses to Smoking in Tobacco Smokers After Overnight Abstinence," *Am J Psychiatry*, vol. 162, pp. 567–577, 2005.

[44] R. A. Wise, "The Neurobiology of Craving: Implications for the Understanding and Treatment of Addiction," *J. Abnorm. Psychol.*, vol. 97, pp. 118–132, 1988.

[45] C. X. Poulos, R. E. Hinson, and S. Siegel, "The role of Pavlovian processes in drug tolerance and dependence: implications for treatment," *Addict. Behav.*, vol. 6, pp. 205–211, 1981.

[46] S. T. Tiffany, E. Singleton, C. A. Haertzen, and J. E. Henningfield, "The development of a cocaine craving questionnaire," *Drug Alcohol Depend.*, vol. 34, no. 1, pp. 19–28, Dec. 1993.

[47] M. N. Smolka, M. Bühler, S. Klein, U. Zimmermann, K. Mann, A. Heinz, and D. F. Braus, "Severity of nicotine dependence modulates cue-induced brain activity in regions involved in motor preparation and imagery.," *Psychopharmacology (Berl).*, vol. 184, no. 3–4, pp. 577–88, Mar. 2006.

[48] F. J. McClernon, F. B. Hiott, J. Liu, A. N. Salley, F. M. Behm, and J. E. Rose, "Selectively reduced responses to smoking cues in amygdala following extinction-based smoking cessation: results of a preliminary functional magnetic resonance imaging study.," *Addict. Biol.*, vol. 12, no. 3–4, pp. 503–12, Sep. 2007.

[49] J. M. Wertz and M. A. Sayette, "Effects of smoking opportunity on attentional bias in smokers.," *Psychol. Addict. Behav.*, vol. 15, no. 3, pp. 268–271, 2001.

[50] R. C. deCharms, "Reading and controlling human brain activation using real-time functional magnetic resonance imaging.," *Trends Cogn. Sci.*, vol. 11, no. 11, pp. 473–81, Nov. 2007.

[51] S. Posse, D. Fitzgerald, K. Gao, U. Habel, D. Rosenberg, G. J. Moore, and F. Schneider, "Real-time fMRI of temporolimbic regions detects amygdala activation during single-trial self-induced sadness," *Neuroimage*, vol. 18, no. 3, pp. 760–768, Mar. 2003.

[52] S. S. Yoo, H. M. O'Leary, T. Fairneny, N. K. Chen, L. P. Panych, H. Park, and F. a Jolesz, "Increasing cortical activity in auditory areas through neurofeedback functional magnetic resonance imaging.," *Neuroreport*, vol. 17, no. 12, pp. 1273–8, Aug. 2006.

[53] S. S. Yoo, J. H. Lee, H. O'Leary, V. Lee, S. E. Choo, and F. A. Jolesz, "Functional magnetic resonance imaging-mediated learning of increased activity in auditory areas," *Neuroreport*, vol. 18, no. 18, pp. 1915–1920, 2007.

[54] R. C. deCharms, F. Maeda, G. H. Glover, D. Ludlow, J. M. Pauly, D. Soneji, J. D. E. Gabrieli, and S. C. Mackey, "Control over brain activation and pain learned by using real-time functional MRI," *PNAS*, vol.102, no. 51, pp. 18626-18631, 2005.

[55] N. Weiskopf, R. Sitaram, O. Josephs, R. Veit, F. Scharnowski, R. Goebel, N. Birbaumer, R. Deichmann, and K. Mathiak, "Real-time functional magnetic resonance imaging: methods and applications.," *Magn. Reson. Imaging*, vol. 25, no. 6, pp. 989–1003, Jul. 2007.

[56] A. Caria, R. Veit, R. Sitaram, M. Lotze, N. Weiskopf, W. Grodd, and N. Birbaumer, "Regulation of anterior insular cortex activity using real-time fMRI.," *Neuroimage*, vol. 35, no. 3, pp. 1238–46, Apr. 2007.

CHAPTER 3

PREDICTION OF SUBJECT'S BRAIN STATES

In recent years, machine learning techniques have gained tremendous popularity in medical image analysis. Temporal brain state classification is one of the major applications of machine learning techniques using functional MRI brain data. Typically, in brain-state classification experiments, the machine learning algorithm is used to build a model based on multiple examples from two different classes. These classes are defined based on the stimulus that is presented to the subject and their corresponding brain state. And the examples are brain volumes captured when the subject is in that particular brain state. Once the model is built, it can be used to predict the brain state of each new unseen brain volume. When used for regression, this model can be used to learn and predict an outcome on a continuous scale instead of binary output in the case of classification.

The major application of brain state classification is to study mental representations and enable brain mapping. This can be implemented for various applications like classifying brain states while craving vs. not craving, activation vs rest or even more complicated experiments like lie detection. In this chapter, we first look at brain-state classification of activation vs rest to compare and contrast three different acquisition techniques. Then, we look at classification of controlled pressure pain vs rest in patients with temporomandibular disorders (TMD) and healthy controls. Lastly, we report findings from our experiment in characterizing graded fMRI activation.

3.1 Support vector machine classification of Arterial Volume-weighted Arterial Spin Tagging (AVAST) images

This section explores the use of support vector machine (SVM) classification technique with motor-visual activation paradigm to perform brain state classification into activation and rest. Images were acquired using a recently developed variant of traditional pseudo-continuous arterial spin labeling (PCASL) technique called arterial volume-weighted arterial spin tagging (AVAST). The classification scheme is also performed on images acquired using blood oxygenation level dependent (BOLD) and traditional perfusion-weighted arterial spin labeling (ASL) techniques for comparison. The results demonstrate that this technique outperforms traditional pseudo-continuous ASL, achieving classification accuracy comparable to that of BOLD contrast. AVAST demonstrates superior signal-to-noise ratio (SNR) and improved temporal resolution as compared to traditional perfusion-weighted ASL, and reduced sensitivity to scanner drift as compared to BOLD. Owing to these characteristics, AVAST lends itself as an ideal choice for dynamic fMRI and real-time neurofeedback experiments with sustained activation periods.

3.1.1 Introduction

Functional magnetic resonance imaging (fMRI) is a non-invasive technique used for visualization of regional brain activity. Blood oxygenation level dependent (BOLD) contrast based techniques are the most commonly used methods for acquiring fMRI images. The hemodynamic response to neuronal activation entails a temporary increase in blood volume and oxygenation level in the blood. BOLD techniques take advantage of the difference in magnetic properties of oxygenated and deoxygenated blood to generate images and are widely used as a marker for providing reliable information about neural activation [1]. The intensity of obtained

images is relative and not individually quantitative because BOLD does not involve direct measurement of any physiological parameter, unless a number of additional measures are collected [2, 3]. The BOLD signal is sensitive to field inhomogeneities caused by the differences in magnetic susceptibility of air and tissue which may results in local tissue distortions and signal losses [4, 5]. The within-session slow scanner drift which has been observed in most studies using BOLD imaging [6] makes it impractical to use for studies involving neural processes with long activation periods.

The observed signal in arterial spin labelling (ASL) method depends on cerebral blood flow (CBF) alone and is independent of the oxygenation level. ASL techniques are less sensitive to local susceptibility artefacts due to the use of shorted echo times (TEs) and show reduced sensitivity to the MR scanner drift since they are subtraction techniques. Unlike BOLD imaging, they are also capable of absolute quantification of CBF in well characterized physiological units, but ASL suffers from inadequacies such as low SNR, and poor temporal resolution.

Since cerebral perfusion is regulated at the arteriolar level, measuring the arterial cerebral blood volume (aCBV) provides useful information about neuronal activation [7,8]. Arterial volume-weighted arterial spin tagging (AVAST) is a variant of pseudo-continuous arterial spin labeling acquisition (PCASL) technique which measures aCBV. fMRI using physiological parameters such as CBF or CBV, unlike BOLD fMRI, provides a quantifiable contrast and is more closely related to neural activity [9, 10, 11].

The short relaxation time of arterial blood causes the tag to decay rapidly resulting in lower SNR while using traditional perfusion-weighted ASL techniques. AVAST demonstrates superior SNR since the images are acquired while tag is still in the arteries before perfusion. AVAST is based on optimizing the timing parameters of a PCASL sequence such that the

subtracted ASL signal is predominantly from the arteries, rather than from the capillaries and the tissue parenchyma. The technique tailors the tagging duration and repetition time (TR) for each subject to achieve a contrast that depends on aCBV with little contamination from the perfusion signal by taking advantage of the kinetics of the tag through that subject's vasculature. AVAST exhibits activation detection sensitivity and temporal resolution which is on par with BOLD imaging and an improvement over the standard CBF ASL technique, while preserving its quantitative nature and statistical advantages [12].

Machine learning techniques have been used increasingly to analyze fMRI images. Machine learning involves the use of an algorithm to facilitate learning from examples. In supervised machine learning algorithms, there is first a training phase, during which labeled input training samples are used to build a model that captures the relationship between the training samples and the corresponding labels [13]. This model is then used during the testing phase to compute an output label for any new testing data sample. Such a setup has been used with functional magnetic resonance imaging (fMRI) data to enable brain-state classification [14, 15, 16].

The objective of this work is to determine whether AVAST is suitable for SVM classification and to compare its performance to BOLD and perfusion-weighted ASL in terms of SVM classification accuracy. To that end, we employed the SVM classification algorithm to perform temporal brain state classification of motor-visual activation vs. rest using images captured by each of the three acquisition techniques : BOLD, ASL and AVAST. We have presented their comparative performance in terms of classification accuracy and significant model weights. Since ASL has a lower temporal resolution, we collect fewer samples in ASL as compared to BOLD and AVAST in the same duration. We have repeated analysis with

subsampled BOLD and AVAST runs and presented results to ensure that the classification was not driven by number of examples.

3.1.2 Methods

3.1.2.1 Stimulation paradigm

Ten healthy subjects participated in this study after signing a written consent and were scanned in accordance with the Institutional Review Board of the University of Michigan. They were given mirrored glasses to view a rear projection screen while being scanned. The paradigm involved displaying 5 cycles of alternating 30s blocks of flashing checkerboard (8Hz) and static fixation cross (total duration = 300 seconds). The subjects performed a robust visuo-spatial activation task and the images used for the stimulation paradigm are shown in figure 3.1. Subjects were instructed to perform self-paced finger tapping with their right hand when presented with the flashing checkerboard, and rest when presented with the fixation cross. The experiment was performed twice per subject and two runs of data were acquired using each acquisition technique.

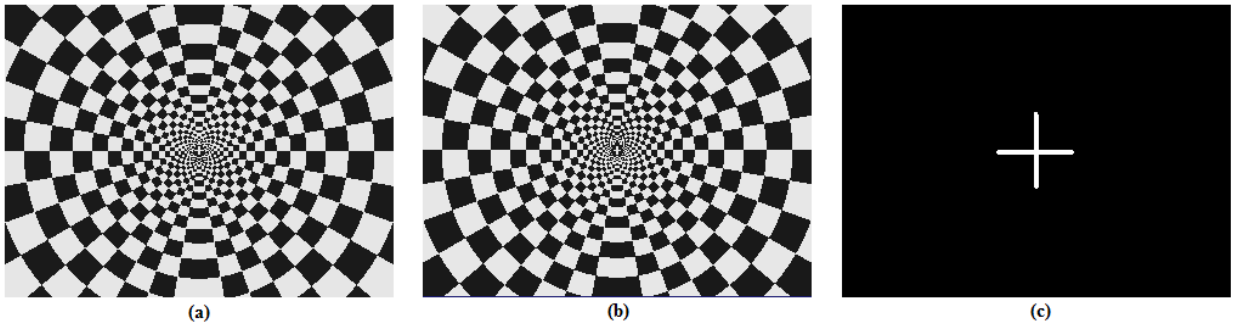

Figure 3.1: Images used for the stimulation paradigm.
(a) Checkerboard, (b) Inverted checkerboard, (c) Fixation cross

3.1.2.2 Data acquisition

All functional images were collected on a 3T GE Signa Excite scanner. Images were captured using each of the three acquisition techniques (BOLD, ASL, AVAST) on every subject while they performed the activation task. To ensure that the steady state was reached, four dummy scans were collected at the start of each run and disregarded. Image acquisition details are as follows:

BOLD: A single-shot gradient echo reverse spiral pulse sequence was used (Scan parameters- TR/TE/FA/FOV=2s/30ms/90°/24cm, 64x64matrix, 22 contiguous slices)

Perfusion-weighted ASL: Images were acquired using a functional CBF scheme employing an off-resonance corrected PCASL technique [17] followed by a 3D stack of spirals acquisition [18]. (Scan parameters- TR/TE/FOV=4s/4.5ms/24cm, tagging duration = 2s, post inversion delay = 1.5s, 64x64matrix, 22 contiguous slices, slab thickness = 11cm, bandwidth = 125KHz)

AVAST: The same pulse sequence utilized for the perfusion-weighted scans was used here, but the tagging parameters were modified to achieve arterial blood weighting. Firstly, a calibration scan was implemented in order to find the optimal timing parameters (tagging duration and TR) for each subject as in [12]. Using these parameters that were tailored for each subject, images were then acquired using the functional aCBV scheme of AVAST (i.e. adjust TR and tagging duration obtained from the calibration scan, no post inversion delay)

3.1.2.3 Preprocessing steps

All datasets were reconstructed and the following preprocessing steps were performed before analysis using the support vector machine training and testing setup.

BOLD: A custom MATLAB code was used for k-space spike removal and spiral reconstruction. SPM8 [19] was used to perform the following: (1) slice timing correction, which corrects for differences in acquisition time between slices during sequential imaging. (2) rigid body motion correction and (3) spatial smoothing using a Gaussian smoothing kernel with FWHM of 8 mm.

ASL and AVAST: As before, custom MATLAB code was used for k-space spike removal and 3D spiral reconstruction and SPM8 motion correction was performed. Next, the resulting time series of images were surround subtracted and spatial smoothing was performed using a Gaussian smoothing kernel with FWHM of 8 mm. Only 11 slices that were consistent across BOLD, ASL and AVAST and covered the motor and visual cortices were used for further analysis.

3.1.2.4 Features and examples

A classifier function outputs a binary class label for every input set of feature values. The features represent an example, whereas, the label signifies the class that a particular example belongs to. More specifically, if x is an example with features [x_1, x_2, x_3,..] and the class label is denoted by y=(±1), then the classifier function f( ) computes the label for every given input, i.e. y=f(x).

In our study, at each time point, a brain activation volume is acquired (using one of the three acquisition techniques). Each such volume is used as a separate example in which the voxel grey scale intensities act as features. Depending on whether the subject was tapping their finger or resting, a label of +1 or -1 is associated with each example.

Data acquired during one of the two runs is used as training data while the other separate run is used as testing data. In the training phase, a mapping is learned from the training examples

to the respective class labels and a classifier is built. In the testing phase, this model is used to predict the class of a previously unseen example from the testing data. Classifier performance is calculated as the ratio of the number of correctly classified test examples to the total number of test examples. At first, run 1 was used for training the classification model, whereas, run 2 was used to assess the effectiveness of the model and then vice versa.

3.1.2.5 Dataset dimensionality

Each of the acquired 3-dimensional volumes, which act as training and testing examples, was of size [64x64x11] voxels. Initially, this accounted to 45,056 features that were then reduced to ~8,000 features by excluding all voxels that fell outside the brain region by using a brain mask created for each individual subject. For BOLD images, this mask was computed by including only those voxels that were within one standard deviation of the mean of the mean image. For ASL and AVAST, all voxels within one standard deviation of the mean of the baseline image (i.e., the mean of the control images in the time series) were included. Note that, for each of the techniques, a different number of examples were collected due to the different TR in each technique (see table 3.1). The number of examples also varies because of the use of surround subtraction in ASL and AVAST. Two such runs were collected per subject.

The runs acquired with perfusion-weighted ASL technique (TR = 4s) include fewer time points than those acquired using BOLD (TR = 2s) or AVAST (TR = 1.9 to 2.5s) In order to address this discrepancy, the analysis was repeated for subsampled runs of BOLD and AVAST such that only every other time point was considered during analysis. These are denoted as sBOLD and sAVAST respectively. The effective TR for sBOLD is 4s and the effective TR in the case of sAVAST ranges from 3.8s to 5s as noted in table 3.1.

Technique	TR	#Examples
BOLD	2s	150
ASL	4s	74
AVAST	1.9s	156
AVAST	2.0s	148
AVAST	2.1s	142
AVAST	2.2s	134
AVAST	2.4s	124
AVAST	2.5s	118
sBOLD	4s*	75
sAVAST	3.8s to 5s*	59 to 78

Table 3.1: Number of examples in train and test datasets corresponding to each TR using each acquisition technique. sBOLD and sAVAST are the subsampled runs of BOLD and AVAST. * denotes effective TR after subsampling

3.1.2.6 SVM classification

In standard SVM classification approach, a separating boundary between the two classes of examples (e.g. +1 and −1) is learned such that the distance (termed "margin") between the data points and boundary is maximal. In higher dimensions, the separating boundary manifests itself as a hyperplane. This separating hyperplane generated by the SVM algorithm is orthogonal to the weight vector **w** which defines the direction in which the examples of the two classes differ most from one another. Thus, the classifier is parameterized by **w,** which can be solved for by using the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{w}} \| \mathbf{w} \|^2 + C \sum_{i=1}^{n} \xi_i \\ &s.t.\ \forall i \in \{1, 2, ..., n\}: \\ &\mathbf{y}_i \mathbf{w}^T \mathbf{x}_i \geq 1 - \xi_i \ \& \ \xi_i \geq 0 \end{aligned} \qquad (1)$$

where $\mathbf{w}$ is the normal vector to the hyperplane, $\mathbf{y}_i$ are the known input class labels, $\mathbf{x}_i$ are the input feature vectors, C is the trade-off parameter used to penalize misclassifications and ξ_i are the non-negative slack variables which measure the degree of misclassification of the input data $\mathbf{x}_i$. The SVM then uses the sign of the decision function $f(x)=\mathbf{w}^T\mathbf{x}$ to classify any new data point x represented by the feature vector $\mathbf{x}$ into one class or the other.

LIBSVM [20], a Library for Support Vector Machines, was used to perform the SVM classification with the default linear kernel and default value of C = 1.

3.1.2.7 Transition periods

The paradigm design includes 10 alternating blocks of flashing checkerboard and fixation cross. Thus, there are 9 intervals when the subject switches from one state to another (finger tapping or rest). These are called transition periods during which the vascular response to neuronal activation is still "ramping up" to its stable state in our block design experiment. We investigated the effect of the transition period by excluding scans acquired during these periods from the modeling exercise as in [21], as follows. Initially, the training was implemented using all the time points. Then this exercise was repeated by excluding 1 time point from both blocks (last time point from previous block and first time point from next block) at each transition point. The same was further repeated by excluding 2 and 3 time points, respectively. This timing is illustrated in the figure 3.2.

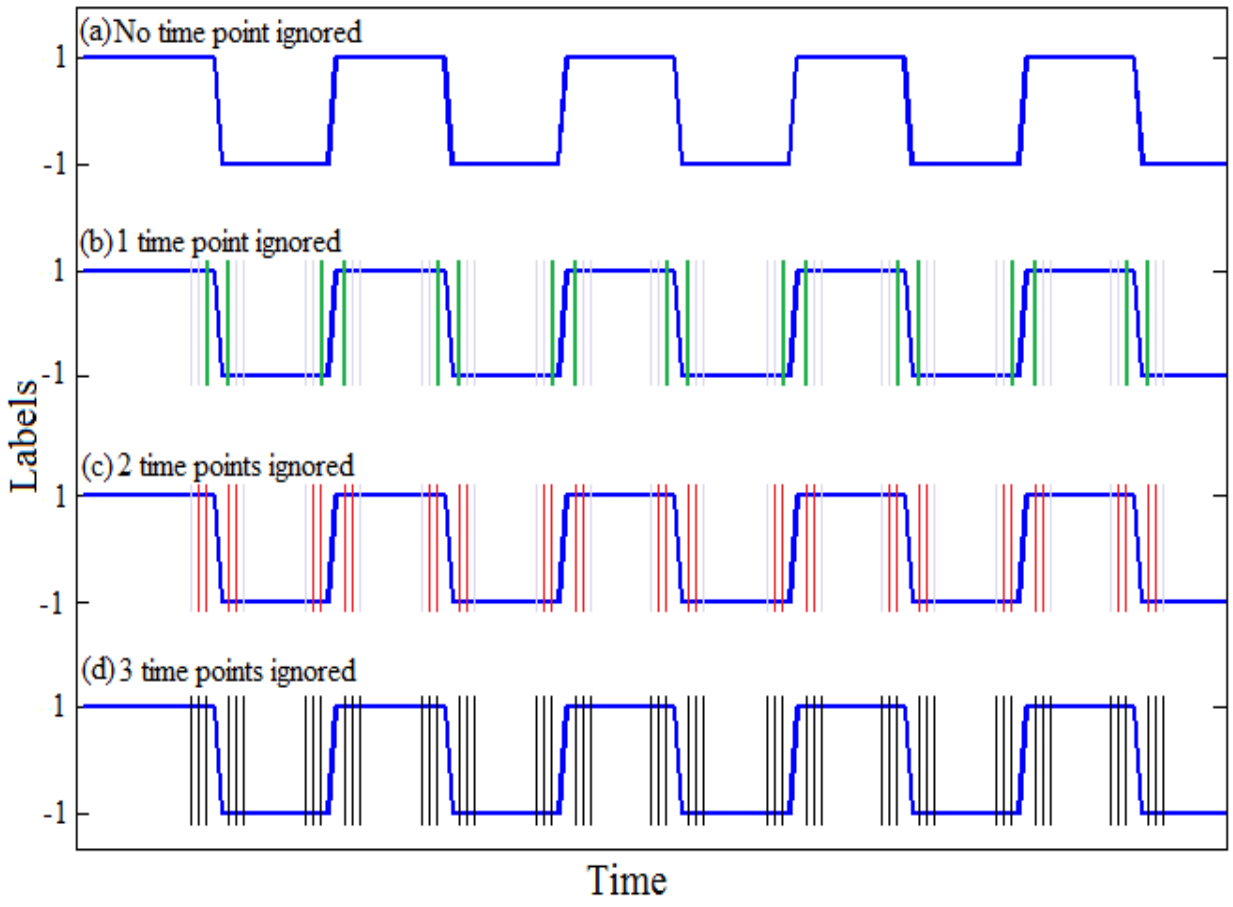

Figure 3.2: Schematic diagram depicting the transition points that are ignored
(a) No time points ignored, (b) 1 time point ignored – green,
(c) 2 time points ignored – red, (d) 3 time points ignored – black

3.1.2.8 Permutation tests

The weight vector **w** generated by the SVM algorithm is representative of the most discriminatory regions of the brain. When mapped back into original image space, this vector generates the discriminating volume also called the weight vector map. SVM is a multivariate pattern analysis technique which assigns a separate weight to each voxel which depends on the weights assigned to other voxels. The weight vector map is thus a representation of the voxels that are most vital to the classification. The magnitude of the absolute value of each voxel weight

determines its relative importance in discriminating the brain states and the most important voxels for discrimination of the brain states can be inspected by thresholding the obtained weight vector map.

A permutation test was employed to assess the reproducibility of these spatial patterns. Briefly, permutation tests are nonparametric techniques that empirically estimate the distribution of a statistic under a null hypothesis empirically and have been used with fMRI data previously in [22, 23, 24]. The null hypothesis proposes that there are no differences between the two brain states and thus the labels assigned to each example are inconsequential. The alternate hypothesis, on the other hand, claims that the assigned class labels are actually indicative of the brain state that an example belongs to and better than random. We can estimate the distribution of weights assigned to each voxel under the null hypothesis by randomly permuting the class labels multiple times and training the SVM each time with this different permutation of labels. In each instance, the weights were normalized to have unit standard deviation. The SVM training was also done once with the known correct non-permuted labels. Now, for each voxel, the p value under null hypothesis was calculated as the ratio of number of times that the voxel weight assigned to it was greater than or equal to the weight assigned to it when training with original non-permuted labels. Since we permuted the labels and trained 2000 different models, if this number is smaller than 20, then that voxel is likely to be predictive of the class label with a significance level of 1%. The weight vector maps shown in the RESULTS display all significant voxels with p value < 0.01.

3.1.3 Results

3.1.3.1 Classification accuracy

Figure 3.3 shows a plot of the mean classification accuracy across both runs of all ten subjects. It demonstrates that the mean classification accuracy obtained by AVAST was consistently better than that offered by ASL and almost equivalent to BOLD. Ignoring transition points improves the classification accuracy initially but plateaus for BOLD and AVAST, whereas, it deteriorates for ASL when 3 time points are ignored in each block.

ASL technique has a lower temporal resolution as compared to BOLD and AVAST. For the same duration of experiment, more images are captured in the other two techniques and thus fewer samples are obtained for analyzing the classifier on ASL images. In order to ensure that the classifier power was not being driven by the number of examples in each case, subsampled BOLD (sBOLD) and subsampled AVAST (sAVAST) runs were also analyzed using the same setup. The mean classification accuracy for these analyses was found to be similar to the earlier case as seen in the plot.

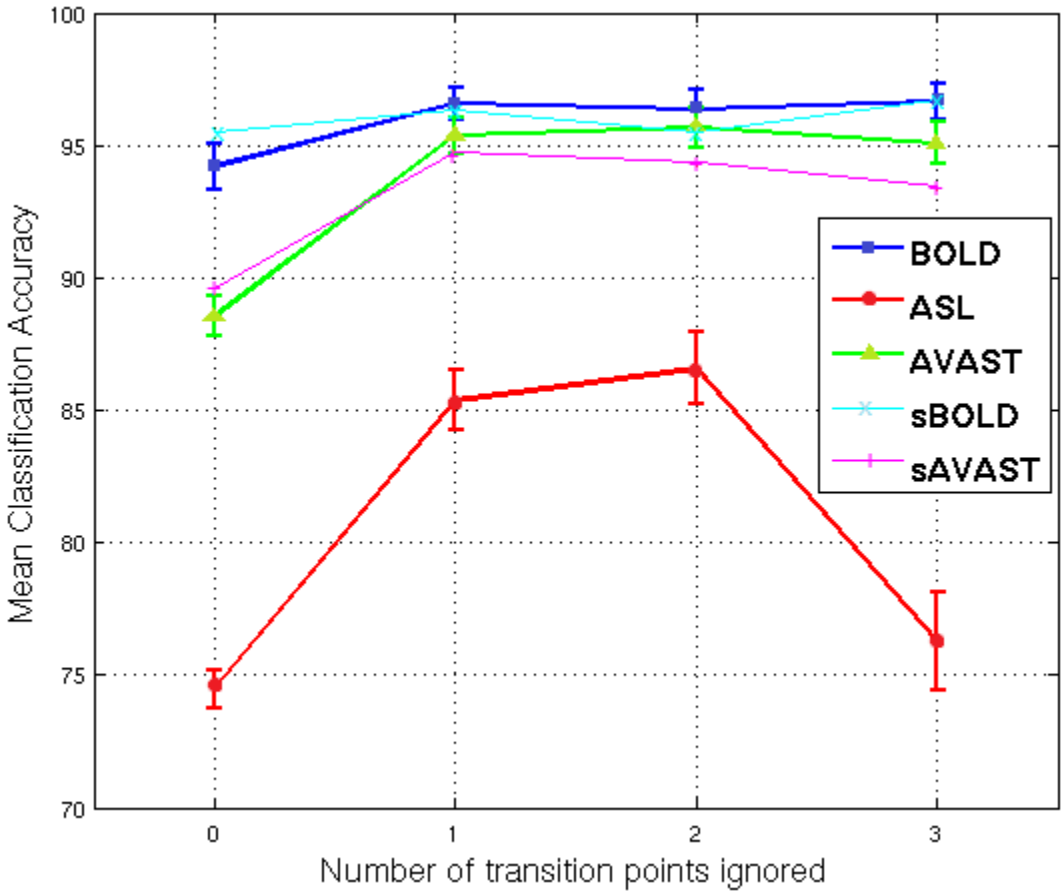

Figure 3.3: Mean classification accuracy over two runs of all ten subjects for each acquisition technique (BOLD-blue, ASL-red, AVAST-green) against number of ignored transition points. The error bars depict standard error.

3.1.3.1 Weight vector maps

The SVM algorithm generates discriminatory weight maps and the permutation tests allow us to find the significant voxels from these maps as described in the METHODS. These maps are indicative of the detection sensitivity of the acquisition technique.

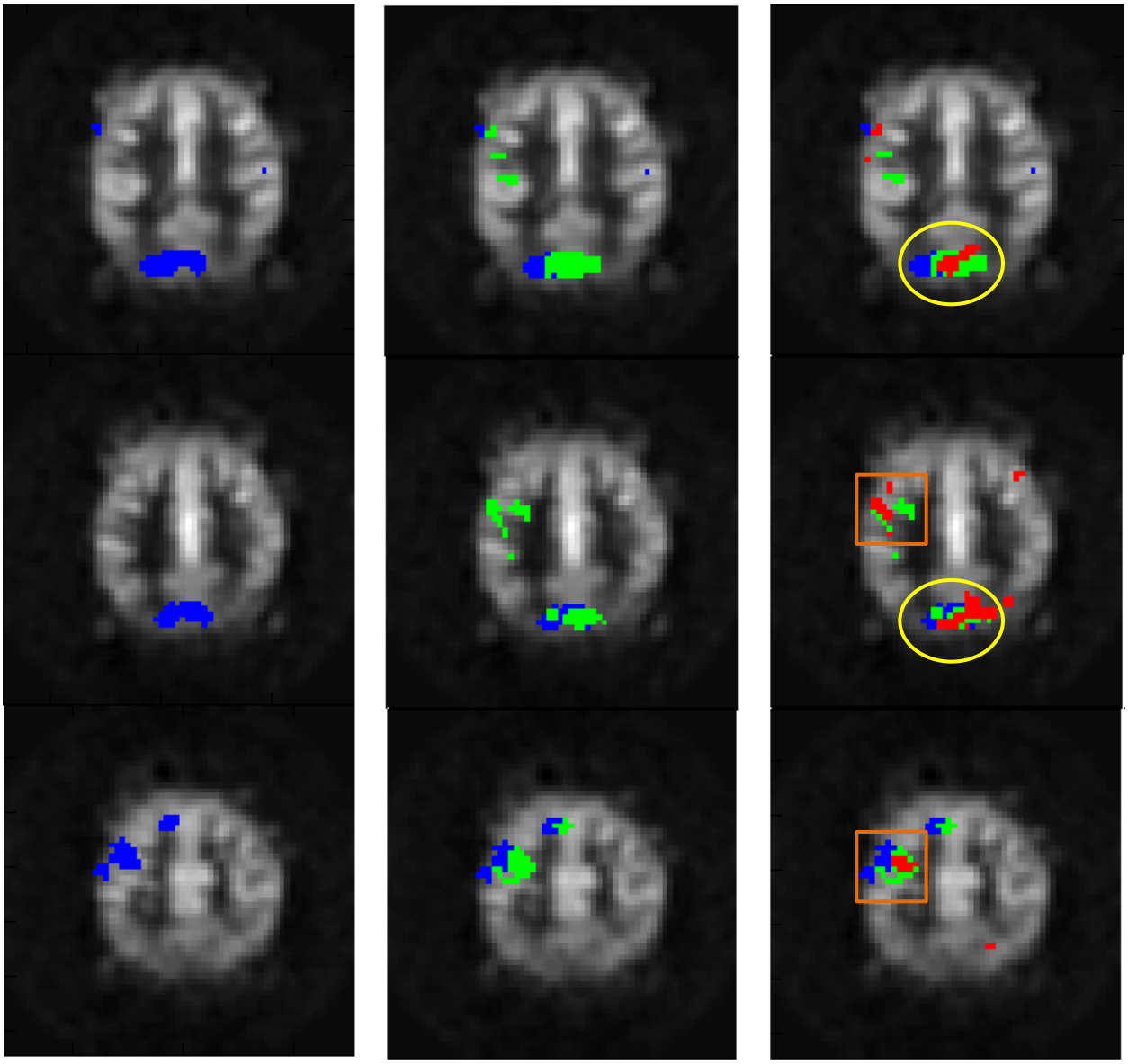

Figure 3.4: Significant SVM weights after permutation tests with $p<0.01$ for a representative subject for each acquisition technique (BOLD-blue, ASL-red, AVAST-green) (a) BOLD, (b) BOLD + AVAST, (c) BOLD + AVAST + ASL
Motor cortex is delineated by orange square and visual cortex is indicated by yellow oval.

Figure 3.4 depicts select slices for a representative subject showing SVM weights in the left motor and premotor cortices (orange squares) and visual cortices (yellow ovals) as expected. The left-most column shows the most significant SVM weights with $p<0.01$ for BOLD (blue) technique. The middle column shows the AVAST (green) weights superimposed on BOLD and the right-most column shows ASL (red) weights superimposed on BOLD and AVAST. The significant weights were bigger and more robust in the AVAST technique as compared to ASL.

3.1.4 Discussion

Traditionally, perfusion-weighted ASL techniques suffer from low SNR and detection sensitivity. By taking advantage of the kinetics of the tag through the vasculature, AVAST facilitates the tailoring of the timing parameters for each subject [12]. This permits the acquisition rate to be much faster and allows much superior temporal resolution as compared to standard perfusion-weighted ASL. Thus, we can acquire a larger number of volumes in the same duration. Also, as noted previously in [12], AVAST offers much better activation detection sensitivity. Both these features are advantageous for the machine learning techniques since it increases the degrees of freedom and also the images obtained are much more sensitive to activation. Thus, AVAST images exhibit better classification performance in terms of higher classification accuracy and more robust clusters of significant weights in the expected brain regions.

The data presented here indicate that AVAST images can be used for SVM classification more reliably than perfusion weighted ASL, and are comparable to BOLD images in terms of their reliability. AVAST images, however, retain some of the advantages of ASL imaging, such as its robustness to scanner drifts and ability to be quantified. ASL images do not depend on T2* contrast, so they can use shorter echo times and thus mitigate susceptibility artifacts.

This study presents promising results that promote the use of machine learning techniques for brain state classification of images acquired by using the AVAST technique. This technique might be used for dynamic fMRI experiments and real-time brain state classification studies as in [25].

3.2 Temporal brain state classification of pain vs rest in healthy control subjects and subjects with temporomandibular disorders

3.2.1 Introduction

Temporomandibular disorders (TMD) are a group of disorders affecting the temporomandibular joints (the joints which connect the jaw bone to the skull) and the muscles of mastication (the muscles that move the jaw) [26]. Pain is the defining feature of TMD and the primary reason for seeking care. Other discomforts may involve restricted jaw function and joint noises. In this study, fMRI scans of healthy control subjects as well as patients with TMD were captured while delivering controlled levels of pain using tools to the face or thumb. Support vector machine (SVM) classification has been used to classify pain vs rest states. The results present significant differences in brain activation related to pain regulation in controls and patients for two separate experiments that involved application of pain stimulus to face or thumb.

3.2.2 Methods

Ten healthy controls and ten TMD patients participated in this study. Each participant was subjected to two 10-minute evoked pressure scans in the MRI scanner and images were collected using a T2*-weighted spiral sequence (TR = 2.5 s, TE = 30 ms, FA = 90°, matrix size 64 x 64 with 48 slices, FOV = 22 cm and 3.44 x 3.44 x 3 mm voxels). Pressure pain was

delivered with a pneumatic system. In the first scan, pain was delivered to the left thumb nail, whereas, in the second scan, pain was delivered to the left anterior temporalis region on the face. Pressures eliciting high and medium pain as previously determined were applied in a pseudo random fashion and interleaved with an "off" condition (no pressure applied). A run contained a total of 12 pain blocks (6 medium, 6 high; each block 25 seconds in duration) and 12 off blocks (each block 25 seconds in duration). The timing of experimental design is illustrated in figure 3.5. Alternate 25s blocks of pain and rest were presented for a total duration of 600 seconds. The medium and high pain blocks were both considered simply as pain blocks in the analysis. SVM analysis was done using LIBSVM [20] package in MATLAB. A model was trained on the first half and then tested on the second half.

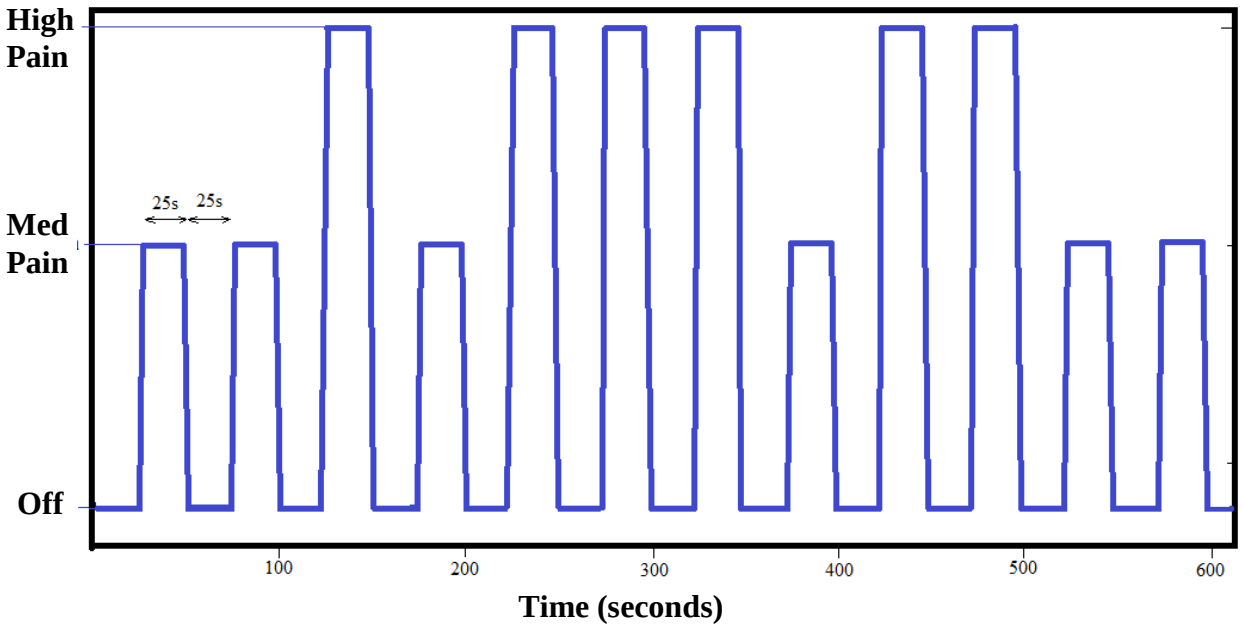

Figure 3.5: Experimental timing for the pain paradigm

In order to ensure the reliability of SVM weights, permutation tests [22, 23, 24] were carried out with 2000 repeats. Thus, a map descriptive of weight significance was generated for each subject called permutation maps. These permutation maps were used to generate the Z maps

for all control subjects and all TMD patients for both thumb as well as face. A t-test was then computed to find differences between the Z maps of all control subjects and all TMD patients with $p<0.01$.

3.2.3 Results

Figures 3.6 and 3.7 show the significant weight differences between control subjects and TMD patients when pain was delivered to the subject's thumb and face respectively. In both cases, significant weights are found in the right insula and the cingulate cortex regions with $p<0.01$. In the case of thumb pain, weights are also observed in the cerebellar region.

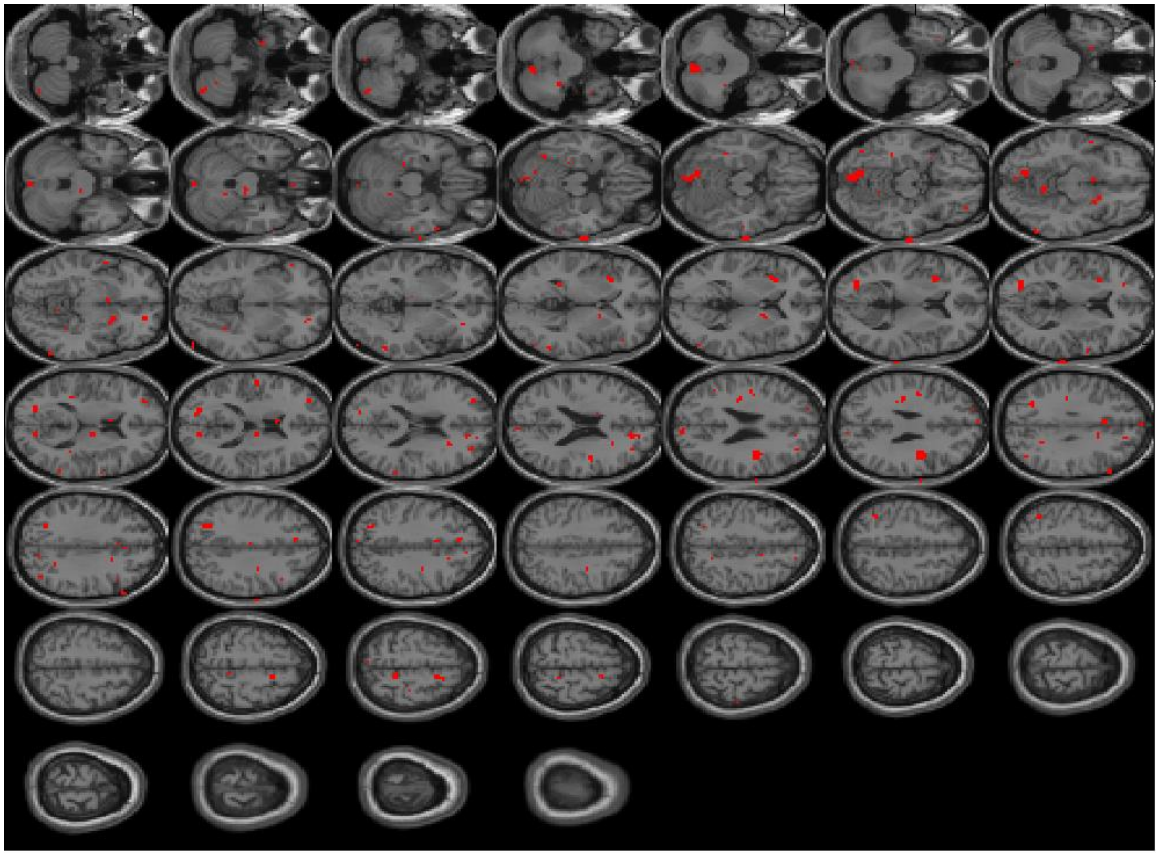

Figure 3.6: Difference map of significant SVM weights between control subjects and TMD patients for Thumb pain with $p<0.01$

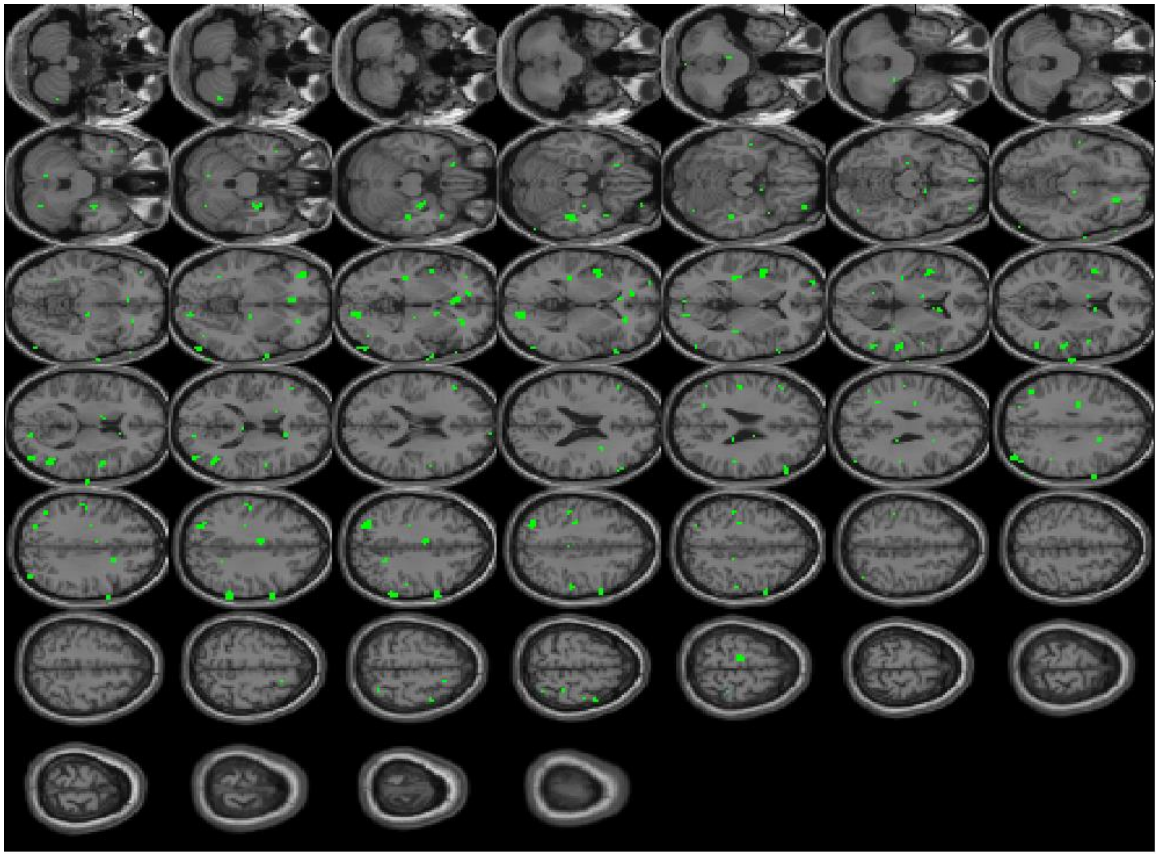

Figure 3.7: Difference map of significant SVM weights between control subjects and TMD patients for Face pain with $p<0.01$

3.2.4 Discussion

It has been shown that patients with TMD show an increased sensitivity to painful stimuli as compared to healthy controls. They have lower thresholds and tolerances for pain and this may be caused due to an alteration in the pain regulatory systems [27]. Thus, it is reasonable that the difference map between patients and controls shows additional significant weights in pain regions of the brain.

The insula is a part of the brain that is involved in consciousness and usually linked to emotion [28]. Insular cortex functions include perception, motor control, cognition and regulation of homeostasis [29]. It has also been associated to the sensation of pain, especially, related to judging the degree of pain [30]. The cingulate cortex is an integral part of the limbic

system which is involved in emotion processing, memory and learning. Neuroimaging has made it possible to link specific aspects of pain perception to localized emotional substrates [31] and thus clearly explain the involvement of cingulate cortex. Although, the exact role of cerebellum in pain processing is unclear, possible functional roles associated with it include perspectives relating to emotion, cognition and motor control in response to pain [32].

In this study, we have introduced another application of SVM based classifier. It has been successfully implemented as a brain-state classifier of pain vs rest. The significant weights observed in the weight map are consistent with literature.

3.3 Characterization of graded fMRI activation using support vector regression

3.3.1 Introduction

When support vector machines are used for classification, then the input training samples are assigned binary labels (typically, ±1) and so the predicted labels for unseen test samples are also binary. These labels correspond to either the brain state or disease state of the scanned subject. When support vector machines are used for regression studies, then the labels assigned to input training samples are on a continuous scale. Thus, the output prediction associated with each unseen test sample is also on a continuous scale and can be used to model and predict a continuous signal [33, 34]. Unlike SVM classification, SVR performs functional approximation using the data and labels from the training run. It then uses the trained model and outputs the value of the function for new unseen test data as described in Chapter 2.

In this section, support vector regression (SVR) has been used in two experiments. In the first, SVR is used to predict graded fMRI activation in the motor cortex. Bilateral finger tapping

at various different frequencies was used for graded motor activation. In the second experiment, SVR is implemented to predict the amount of craving a subject experiences while being presented with smoking stimuli during their scan. The subjects' self-reported craving measures are used as train the model and test the prediction.

3.3.2 Methods

For the finger tapping experiment, the subject used mirrored glasses while being scanned to look at a rear projection screen on which images shown in figure 3.8 were presented. The subjects were instructed to perform bimanual finger tapping at a pace influenced by the images. In each block, the images would transition at 1, 2, 3, or 4Hz. Each such 20s tapping block was interspersed with 20s blocks of static fixation during which the subject was instructed to rest. T_2*-weighted images were acquired on a 3T GE scanner using a custom spiral-in sequence with following parameters (TR/TE/FA/FOV = 2s/30ms/90°/22cm, 64x64 matrix, 40 contiguous axial slices of 3 mm slice thickness). Two such runs were collected per subject. The tapping frequency was used as a label for all fMRI brain volumes acquired within that block. LIBSVM [20] package in MATLAB was used with default parameters to perform SVR. At first, run 1 was used to train a model which was tested on run 2 and then vice versa. Squared correlation coefficient was computed in each case.

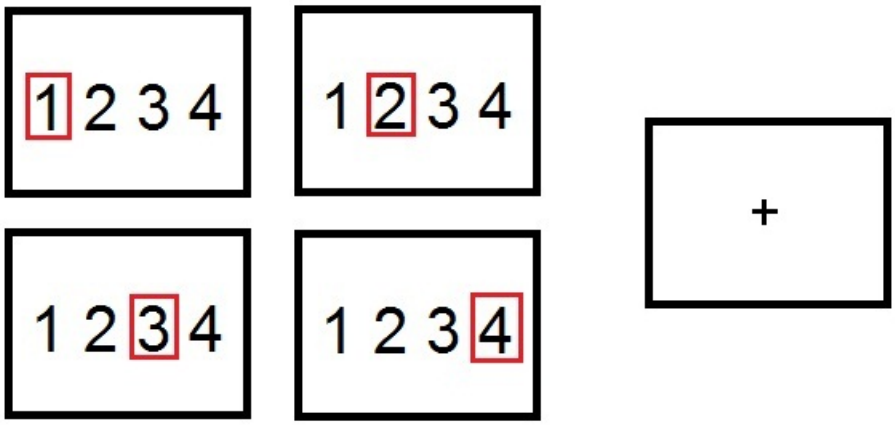

Figure 3.8: Images used to influence the subjects' bimanual finger tapping pace. The numbered images would sequentially transition at 1, 2, 3 or 4Hz to influence tapping frequency for a total of 20s, followed by the fixation cross for same duration.

For the craving experiment, 19 nicotine dependent subjects were scanned. They were presented with a series of images that depicted smoking cues followed by a block of matched set of neutral images devoid of smoking cues. Few examples are demonstrated in figure 3.9. Again, T_2*-weighted images were acquired on a 3T GE scanner using a custom spiral-in sequence with following parameters (TR/TE/FA/FOV = 2s/30ms/90°/22cm, 64x64 matrix, 40 contiguous axial slices of 3 mm slice thickness). Subjects used a MR-compatible button-response pad to rate each image on a scale of 1 to 5 in real-time depending on how much it made them crave cigarettes. 20s blocks of smoking or non-smoking images were followed by a 4s fixation cross for a total scan time of 384 seconds. Again two runs were acquired and LIBSVM [20] package in MATLAB was used to train an SVR model on 1 of the runs and test on the other. Every instance at which the user increased or decreased their craving rating was investigated and that particular instance was accepted as correctly classified only if there was a corresponding change in SVR prediction.

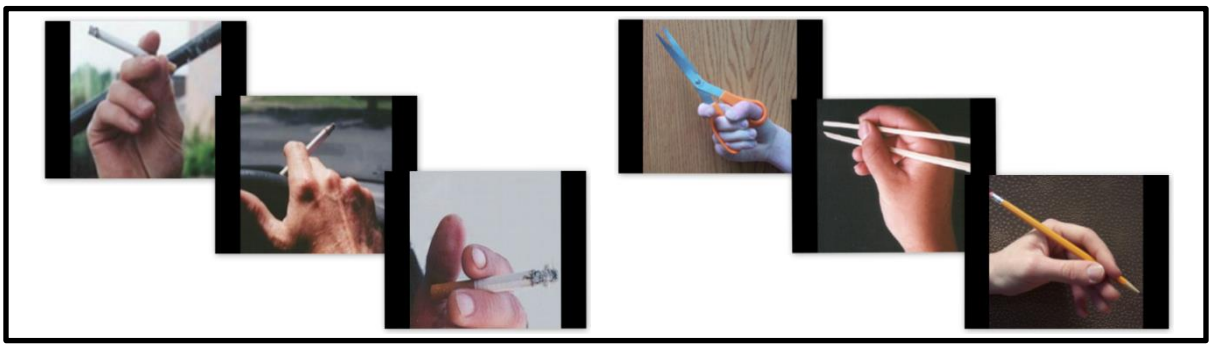

Figure 3.9: Example images showing smoking cues as well as neutral cues. Each image was presented for 4s in 20s blocks and rated by the subject.

3.3.3 Results

Figures 3.10 and 3.11 show, for the finger tapping experiment, the output of SVR prediction (blue) and the expected test labels (red) for a representative subject. The peaks in the SVR output are a clear indication that the algorithm is able to model the graded activation. The first figure shows the output when the model is trained on run 1 and tested on run 2, whereas, the second figure shows the output of SVR when it was trained on run 2 and then tested on new unseen run 1.

The squared correlation coefficient in the first case is 0.69, whereas, in the second case, it is computed as 0.75.

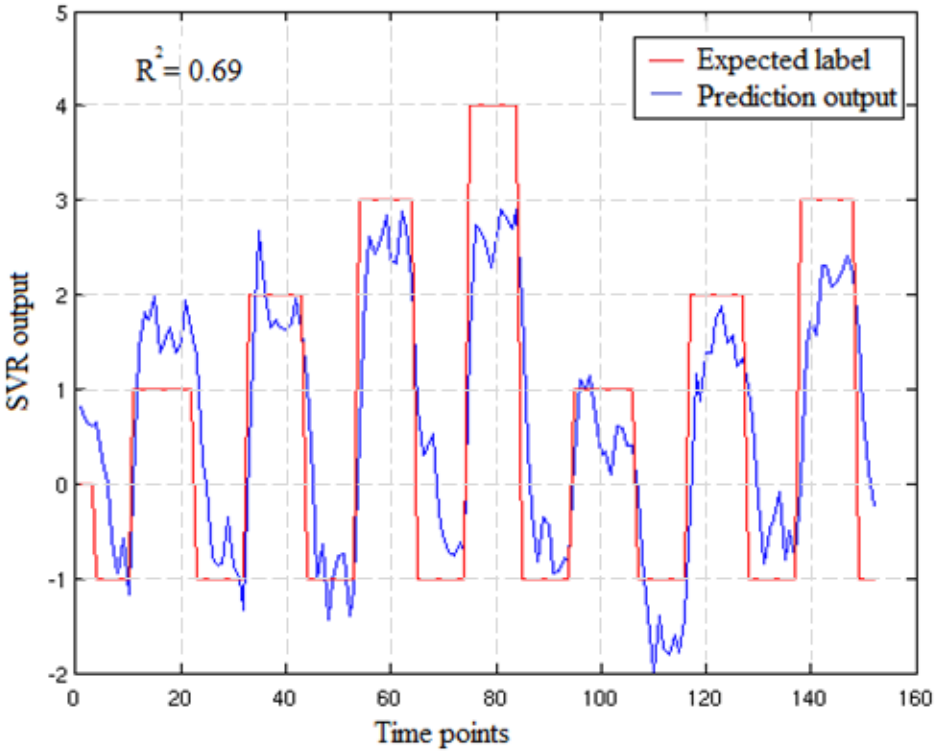

Figure 3.10: SVR output after training on run 1 and testing on run 2 for a representative subject for motor activation paradigm.

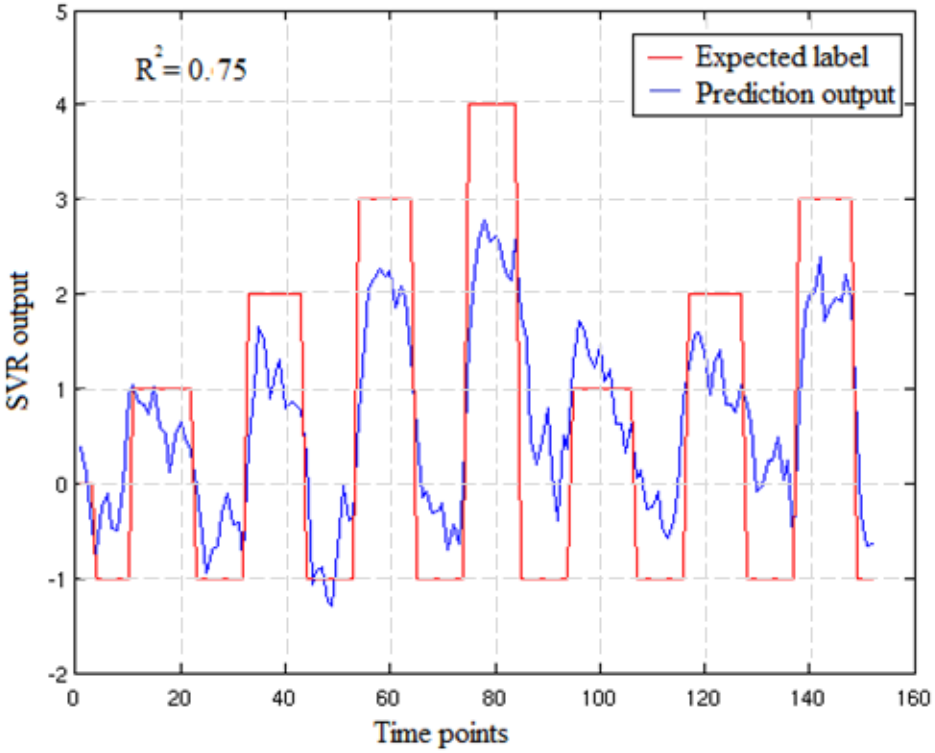

Figure 3.11: SVR output after training on run 2 and testing on run 1 for a representative subject for motor activation paradigm.

Figure 3.12 shows, for the craving experiment, an example plot of SVR prediction (blue) and subject's self-reported craving measure. As per the classification scheme used, the average classification accuracy over 19 subjects was 62%. It can be seen from the output, that SVR is able to predict some of the sudden blips and crests in the subject's self-reported ratings.

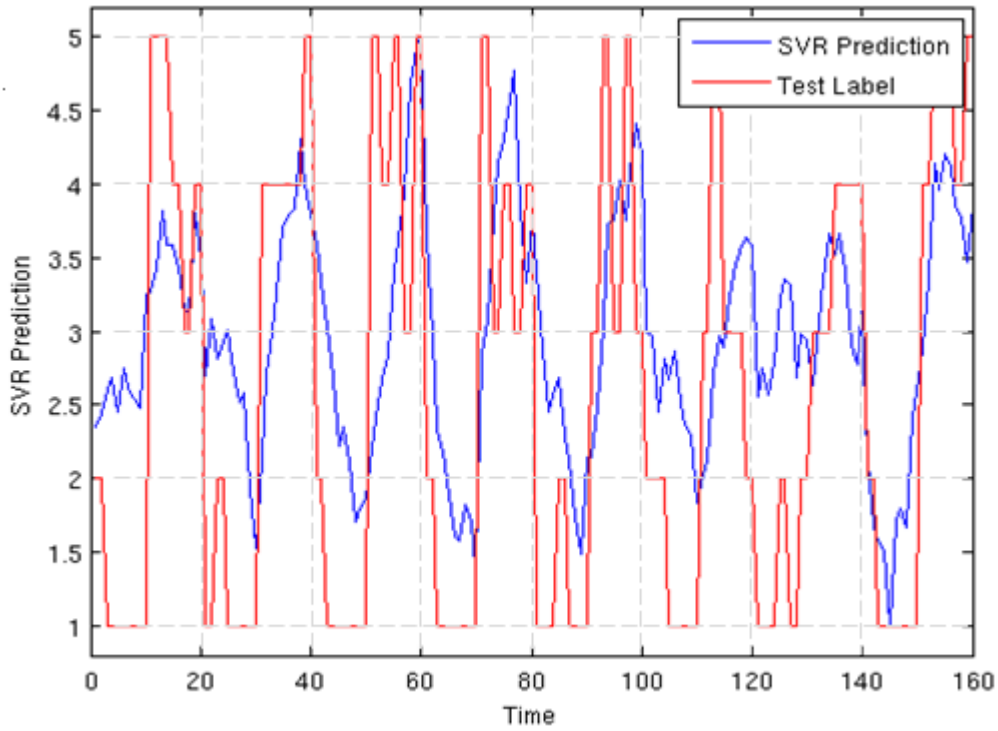

Figure 3.12: SVR output after training on run 1 and testing on run 2 for a representative subject for craving paradigm.

3.3.4 Discussion

Support vector regression has been successfully used to learn from graded fMRI activation and then predict the same for the test run. In the first experiment, this algorithm has been used to model robust motor activation. This can be extended to study other neural processes, especially, those that involve activations that are continuous in nature and can involve gradations. In the second experiment, this has been used to model the amount of craving that a subject is experiencing in real-time. The problem faced here is that there is no known reliable way of measuring the amount of craving externally without relying on a subject's self-reported

craving levels. Hence, it is not easy to quantify the input training labels that are used to build the model. Regardless, the SVR output seems to detect the sudden changes in subjects' self-reported craving levels efficiently in lots of instances.

3.4 Concluding remarks

In this chapter, we saw results from experiments in which machine learning algorithms have been implemented for brain state prediction. This was used to compare acquisition techniques, find brain state differences between healthy subjects and patients and also to investigate graded fMRI activation. In the next chapter, we report findings from our experiment in which machine learning techniques are used for subject classification.

REFERENCES:

[1] J. A. Detre and J. Wang, "Technical aspects and utility of fMRI using BOLD and ASL," *Clin. Neurophysiol.*, vol. 113, no. 5, pp. 621–634, May 2002.

[2] T. L. Davis, K. K. Kwong, R. M. Weisskoff, and B. R. Rosen, "Calibrated functional MRI : Mapping the dynamics of oxidative metabolism," *Proc. Natl. Acad. Sci*, vol. 95, no. February, pp. 1834–1839, 1998.

[3] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike, "Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex," *Proc. Natl. Acad. Sci*, vol. 96, no. August, pp. 9403–9408, 1999.

[4] B. P. Sutton, S. Member, D. C. Noll, J. A. Fessler, and S. Member, "Fast , Iterative Image Reconstruction for MRI in the Presence of Field Inhomogeneities," *IEEE transactions on medical imaging*, vol. 22, no. 2, pp. 178–188, 2003.

[5] N. Weiskopf, C. Hutton, O. Josephs, and R. Deichmann, "Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T.," *Neuroimage*, vol. 33, no. 2, pp. 493–504, Nov. 2006.

[6] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. Yang, J. H. Duyn, and J. A. Frank, "Investigation of Low Frequency Drift in fMRI Signal," *Neuroimage*, vol. 9, pp. 526–533, 1999.

[7] S. P. Lee, T. Q. Duong, G. Yang, C. Iadecola, and S. G. Kim, "Relative changes of cerebral arterial and venous blood volumes during increased cerebral blood flow: implications for BOLD fMRI.," *Magn. Reson. Med.*, vol. 45, no. 5, pp. 791–800, May 2001.

[8] J. A. Mumford and T. Nichols, "Simple group fMRI modeling and inference.," *Neuroimage*, vol. 47, no. 4, pp. 1469–75, Oct. 2009.

[9] T. Jin and S. G. Kim, "Improved cortical-layer specificity of vascular space occupancy fMRI with slab inversion relative to spin-echo BOLD at 9.4 T.," *Neuroimage*, vol. 40, no. 1, pp. 59–67, Mar. 2008.

[10] T. Q. Duong, E. Yacoub, G. Adriany, X. Hu, K. Ugurbil, J. T. Vaughan, H. Merkle, and S. G. Kim, "High-resolution, spin-echo BOLD, and CBF fMRI at 4 and 7 T.," *Magn. Reson. Med.*, vol. 48, no. 4, pp. 589–93, Oct. 2002.

[11] T. Kim and S. G. Kim, "Cortical layer-dependent arterial blood volume changes: Improved spatial specificity relative to BOLD fMRI," *Neuroimage*, vol. 49, no. 2, pp. 1340–1349, 2010.

[12] H. Jahanian, S. Peltier, D. C. Noll, and L. Hernandez-Garcia, “Arterial cerebral blood volume-weighted functional MRI using pseudocontinuous arterial spin tagging (AVAST).,” *Magn. Reson. Med.*, vol. 00, pp. 1–12, Apr. 2014.

[13] F. Pereira, T. Mitchell, and M. Botvinick, “Machine learning classifiers and fMRI: a tutorial overview.,” *Neuroimage*, vol. 45, no. 1 Suppl, pp. S199–209, Mar. 2009.

[14] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[15] T. M. Mitchell, R. Hutchinson, R. S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman, “Learning to Decode Cognitive States from Brain Images,” *Mach. Learn.*, vol. 57, no. 1/2, pp. 145–175, Oct. 2004.

[16] S. M. LaConte, S. Strother, V. Cherkassky, J. Anderson, and X. Hu, “Support vector machines for temporal classification of block design fMRI data.,” *Neuroimage*, vol. 26, no. 2, pp. 317–29, Jun. 2005.

[17] H. Jahanian, D. C. Noll, and L. Hernandez-Garcia, “B0 field inhomogeneity considerations in pseudo-continuous arterial spin labeling (pCASL): effects on tagging efficiency and correction strategy.,” *NMR Biomed.*, vol. 24, no. 10, pp. 1202–9, Dec. 2011.

[18] J. F. Nielsen and L. Hernandez-Garcia, “Functional perfusion imaging using pseudocontinuous arterial spin labeling with low-flip-angle segmented 3D spiral readouts,” *Magn. Reson. Med.*, vol. 69, no. 2, pp. 382–390, 2012

[19] Wellcome Trust Centre for Neuroimaging, “Statistical Parametric Mapping http://www.fil.ion.ucl.ac.uk/spm.” .

[20] C. C. Chang and C. J. Lin, “Libsvm : A library for support vector machines,” *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[21] S. M. LaConte, J. Anderson, S. Muley, J. Ashe, S. Frutiger, K. Rehm, L. K. Hansen, E. Yacoub, X. Hu, D. Rottenberg, and S. Strother, “The Evaluation of Preprocessing Choices in Single-Subject BOLD fMRI Using NPAIRS Performance Metrics,” *Neuroimage*, vol. 18, no. 1, pp. 10–27, Jan. 2003.

[22] T. E. Nichols and A. P. Holmes, “Nonparametric Permutation Tests For Functional Neuroimaging : A Primer with Examples,” *Human Brain Mapping*, vol. 15, pp. 1-25, 2001.

[23] J. Mourão-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, “Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data.,” *Neuroimage*, vol. 28, no. 4, pp. 980–95, Dec. 2005.

[24] H. Jahanian, G.A. Hossein-Zadeh, H. Soltanian-Zadeh, and B. A. Ardekani, “Controlling the false positive rate in fuzzy clustering using randomization: application to fMRI activation detection.,” *Magn. Reson. Imaging*, vol. 22, no. 5, pp. 631–8, Jul. 2004.

[25] L. Hernandez-Garcia, H. Jahanian, M. K. Greenwald, J.K. Zubieta, and S. J. Peltier, “Real-time functional MRI using pseudo-continuous arterial spin labeling.,” *Magn. Reson. Med.*, vol. 65, no. 6, pp. 1570–7, Jun. 2011.

[26] H. R. Mujakperuo, M. Watson, R. Morrison, and T. V Macfarlane, “Pharmacological interventions for pain in patients with temporomandibular disorders,” *The Cochrane Collab.*, no. 10, 2010.

[27] W. Maixner, R. Fillingim, S. Kincaid, A. Sigurdsson, and M. B. Harris, “Relationship Between Pain Sensitivity and Resting Arterial Blood Pressure in Patients With Painful Temporomandibular Disorders,” *Psychosom. Med.*, vol. 59, no. 5, pp. 503–511, 1997.

[28] A. D. Craig, “How do you feel - now? The anterior insula and human awareness,” *Nat Rev Neurosci.*, vol. 10, no. 1, pp. 59–70, 2009.

[29] H. D. Critchley, “Neural mechanisms of autonomic, affective, and cognitive integration,” *J. Comp. Neurol.*, vol. 493, no. 1, pp. 154–66, Dec. 2005.

[30] M. N. Baliki, P. Y. Geha, and A. V. Apkarian, “Parsing pain perception between nociceptive representation and magnitude estimation.,” *J. Neurophysiol.*, vol. 101, no. 2, pp. 875–87, Feb. 2009.

[31] B. A. Vogt, “Pain and Emotion interactions in subregions of the cingulate cortex,” *Nat Rev Neurosci.*, vol. 6, no. 7, pp. 533–544, 2005.

[32] E. A. Moulton, J. D. Schmahmann, L. Becerra, and D. Borsook, “The cerebellum and pain: passive integrator or active participator?,” *Brain Res. Rev.*, vol. 65, no. 1, pp. 14–27, Oct. 2010.

[33] A. J. Smola and B. Scholkopf, “A Tutorial on Support Vector Regression,” http://alex.smola.org/papers/2003/SmoSch03b.pdf - 2003.

[34] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag NY, 2000.

CHAPTER 4

CATEGORIZATION OF SUBJECTS INTO HEALTHY CONTROLS AND PATIENTS WITH DISORDERS

Machine learning techniques can be used to facilitate subject categorization based on their disease state. In such a setting, each subject's data is represented by a single map and the classifier builds a model which learns to classify all maps of control subjects from maps belonging to disordered patients. Thus, the model highlights those parts of the brain that are most discriminatory between the two classes and throws light on the neural correlates of the disorder. In this chapter, we present findings from our study involving categorization of healthy control subjects from patients with Asperger's disorder. Resting state functional connectivity maps are utilized in this study to represent each subject.

4.1 Functional connectivity

Functional connectivity is a statistical concept that is defined as temporal correlation between spatially remote neurophysiological events [1]. It captures statistical dependence between all the different voxels of the brain and is usually computed by finding the temporal correlation between their time courses or other multivariate analysis methods. In fMRI data, low frequency time course fluctuations are found to be temporally correlated between functionally related areas and exist in a number of brain networks. It has been speculated that this interregional correlation in fluctuations is caused by spontaneous neuronal activity and corresponding alterations in blood flow and oxygenation [2]. Functional connectivity cannot give

any information about the directionality of the connection. Recently, most functional connectivity studies are carried out by examining interregional correlations in BOLD fMRI data of subjects in the resting state.

Functional connectivity applies to both task-activated as well as resting-state studies. In task-activated studies, it can refer to correlations across subjects, runs, blocks, trials or individual time points. On the other hand, in resting-state studies, it is assessed across individual BOLD time points during resting conditions. Functional connectivity between brain regions has been estimated primarily using one of two methods: (a) seed-based temporal correlation [3] or (b) spatial independent component analysis (ICA) [4]. The network maps derived using both have been shown to be similar but not identical [5].

4.2 Resting state

An individual is said to be in a resting state when they are not focused on the outside world and their brain is at wakeful rest. The default mode network is a specific network of brain regions that are active when the individual is in this resting state i.e. they show synchronous activations in the absence of any external stimulus and it deactivates in response to any externally demanding goal-directed task [6]. It includes the posterior cingulate cortex, medial prefrontal cortex and lateral parietal cortex. These are regions usually associated with the social cognition network. Thus, deficits in the functional connectivity of these regions imply an impaired or atypical social development. Resting state functional connectivity (RSFC) has gained a lot of interest in recent times to understand the brain's functional organization and to examine if it is altered in any specific group of subjects. These different groups of subjects might be separated based on age, gender or diagnosis of a neurological or psychiatric disorder [7]. RSFC maps were used in [8] to distinguish patients with depression from healthy controls.

4.3 Autism Spectrum Disorders

An autism spectrum disorder (ASD) is a heterogenous, behaviorally defined neurodevelopmental disorder [9]. It is diagnosed on the basis of a triad of behavioral impairments, namely, impaired social interaction, impaired communication and restricted, repetitive interests and activities. The behavioral phenotype associated with ASD is well described but its etiology is not completely understood. Many possible causes of origin have been postulated, namely, genetic, infectious, neurologic, metabolic, immunologic, environmental and other factors. Recent neuroimaging studies have highlighted that it is accompanied by subtle and spatially distributed differences in brain anatomy, mainly brain regions associated with the social cognition network.

4.4 Functional MRI detection of Asperger's Disorder using SVM classification

Asperger's Disorder is a type of high functioning autism, with core deficits in social interaction [9]. Several research studies have been conducted to find distinctive abnormal features from the brain of Asperger's Disorder subjects to find neural correlates of these behavioral characteristics. Recent resting state functional MRI studies have reported that subjects with Asperger's Disorder exhibit decreased functional connectivity between nodes in the default mode network, which is a set of brain regions showing synchronous activations during resting state for healthy subjects. In this study, we employ a machine learning algorithm using support vector machine to differentiate Asperger's Disorder subjects from healthy control subjects based on their resting state functional MRI images. We believe that our method can help identify Asperger's Disorder in an objective and automated manner and serve as a basis for future work to quantify severity of Autism based on resting state MR brain scans.

4.4.1 Introduction

Asperger's disorder is considered one of the autism spectrum disorders and is associated with a history of atypical social development and the presence of repetitive or highly circumscribed behaviors and interests [9]. To find abnormalities of brain structures and functions related to these behavioral characteristics, many functional neuroimaging studies have been conducted [10 - 16]. Particularly, several studies reported that they could successfully identify distinctive brain activation patterns from subjects with Asperger's Disorder using functional Magnetic Resonance Imaging (fMRI) [11].

However, these experiments required that the subjects actively respond to the provided stimuli during scanning, which is potentially problematic with patient populations, as opposed to a resting state study which does not require the use of explicit stimuli. Individuals are said to be in the resting state when they are not focused on the outside world and their brain is at wakeful rest. A resting state functional connectivity (RSFC) study involves detecting temporal correlations in spontaneous blood oxygen level dependent (BOLD) signal oscillations while subjects rest in the scanner.

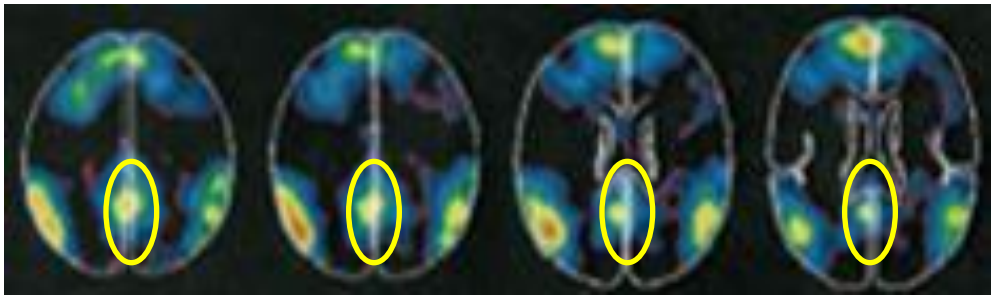

Figure 4.1: Default mode network example showing PCC (yellow oval) [6]

A set of brain regions, collectively termed as the default mode network (DMN), show synchronous activations during a resting state [6]. The DMN mainly comprises of posterior cingulate cortex, medial prefrontal cortex and lateral parietal cortex. It was recently examined in

order to investigate whether there is a significant difference between subjects diagnosed with autism spectrum disorders and healthy control subjects [17].

Functional MRI scans of a resting state brain were used in [17] to form a time series of fMRI images in order to measure the default network connectivity of brain regions. A correlation was computed between time series of each voxel and that of a seed voxel taken from posterior cingulate cortex (PCC), which is a member of the default network. A correlation value above a statistically determined threshold was regarded as activation (or connection between the two nodes) and for Asperger's Disorder subjects, they observed very weak connectivity between the brain regions belonging to the DMN including posterior cingulate cortex, medial prefrontal cortex and lateral parietal cortex as compared to control subjects. Although this method was effective in demonstrating the difference in functional connectivity between subjects with Asperger's Disorder and healthy control subjects, it is subjective to the choice of statistical parameters and it also involved manual visual investigation of the computed correlation maps to decide whether or not the map belonged to a subject with Asperger's Disorder.

In this study, we employ a machine learning approach which involves the training of a support vector machine using the computed correlation map in order to automate this classification in a more objective manner. In typical fMRI applications to identify pathology, the machine learning algorithms learn a functional relationship between features of the brain images and diagnosed state of a subject expressed in terms of a label, which may be assigned discrete or continuous values. This learned functional relationship is used to form a model which is then used to predict the unseen labels for a new test data set. Thus, they facilitate a classifier based predictive learning framework.

Support Vector Machine (SVM) is a powerful set of such machine learning methods. The SVM algorithm seeks a maximum margin separating hyperplane thus making it resilient to overfitting. This means that they provide better generalization, allowing most favorable classification of previously unseen test data [18, 19]. Additionally, when a linear kernel is used, they allow the possibility of generating discrimination maps [20], which describes the contribution of individual training features to classification. Thus, SVMs help not only in effective pattern discrimination but also pattern localization. In our specific application, this feature would help us localize the brain regions that differentiate subjects with the disorder from healthy controls. SVMs can be implemented for use in binary classification as well as continuous regression analysis. Thus, they can be used for differentiating patients from controls as also for quantifying the level of autism severity.

In this study, support vector machines have been implemented to classify resting state functional connectivity data of subjects with Asperger's disorder from that of healthy controls. Preprocessing steps including the use of absolute values, spatial smoothing and a mask that only includes default mode network areas were examined in an attempt to optimize the prediction accuracy of SVM algorithm.

4.4.2 Methods

4.4.2.1 Subjects

Eight adults with Asperger's Disorder (AD) and eighteen healthy control (HC) subjects participated in this study (AD: 20.8±3.2 years, HC: 24.8±5.3 years). The subjects have near or above average intelligence as measured by a standardized IQ test, i.e. the Wechsler Abbreviated Scale of Intelligence (WASI) [21] (AD : mean IQ = 113.4 ± 16.4, HC : mean IQ = 123.1 ± 7.0). Within the patient group, 6 were male and 2 were female; 6 were Caucasian and 2 were Asian-

American. The diagnosis of Asperger's Disorder was based on a clinical assessment completed by a clinical psychologist. The assessment battery consisted of the Autism Diagnostic Observation Schedule-Generic [22], a clinical chart review, and when possible (n = 5), a parent interview (i.e., the Autism Diagnostic Interview-Revised; [23]). All participants met DSM-IV criteria [9] for Asperger's Disorder and had a history of social difficulties beginning in childhood. Ten out of the total eighteen control subjects completed the Young Adult Self Report [24] to rule out any psychiatric symptoms. Remaining eight subjects (27.4 ± 5.3 yrs; 6 male, 2 female) were given a short questionnaire, developed by the authors, to rule out symptoms of inattention, depression, and/or anxiety. Written informed consent was obtained from all subjects prior to enrollment in the study. The study was carried out with the approval of the Internal Review Board of Emory University.

4.4.2.2 fMRI scans

Resting state MR scans were acquired on a 3T Siemens Trio scanner using an echo-planar imaging sequence with TR/TE/FA/FOV of 750 ms/35 ms/50°/22 cm. Ten contiguous 5-mm thick axial slices were acquired in each TR, with an in-plane resolution of 3.44 mm x 3.44 mm, covering the anterior, middle and posterior cingulate cortex and the dorsomedial prefrontal cortex. Resting state data were acquired while subjects were inactive (lying still with visual fixation cross). A total of 280 volumes were collected during a 3.5 minute scan.

4.4.2.3 Functional connectivity map generation

For all individual data, a transformation into MNI space was calculated using SPM2 [25]. The coordinates of a seed voxel in the posterior cingulate cortex were identified in the standard MNI space (MNI: -2, -51, 17; BA 23) as in [26], and then transformed to the original (native) acquisition space. The fMRI signal timecourses for this voxel and four neighboring voxels were

averaged together, and low-pass filtered (< 0.08 Hz) to avoid artifacts (e.g. cardiac and respiration effects), while preserving frequencies contributing to functional connectivity [27, 28]. This timecourse was then used as the reference timecourse. The timecourses of every other voxel in the entire imaged volume were also low-pass filtered. These timecourses were then correlated with the reference timecourse to form functional connectivity maps (i.e., low frequency timecourse correlation maps). The resultant functional connectivity correlation maps were then transformed to MNI space using the transformation computed earlier.

4.4.2.4 Preprocessing steps

Ten axial slices were initially acquired for each subject in their native space. However, due to different head positioning of individual subjects, after spatial normalization only four common slices in MNI space covering the default mode network (ranging from z = +10 mm to z = +30 mm) were used in the subsequent analysis. Also the data from two subjects could not be included in the analysis because it was visually observed to have abnormal transformations. Using the data for SVM training without such initial filtering would definitely deteriorate the classification performance and hence some of the data had to be disregarded. This initial filtering process finally resulted in seven Asperger's Disorder subjects and seventeen control subjects.

The data initially used to train the SVM were correlation values that were normalized using Fisher z-transformation, which is defined by the transformation

$$z = \frac{1}{2}\ln\frac{1+r}{1-r} \qquad (1)$$

where r is the correlation coefficient. Thus z-maps are formed for each subject. Figure 4.2 shows example z-maps of two control subjects and two Asperger's Disorder subjects, representative of the difference in their functional connectivity.

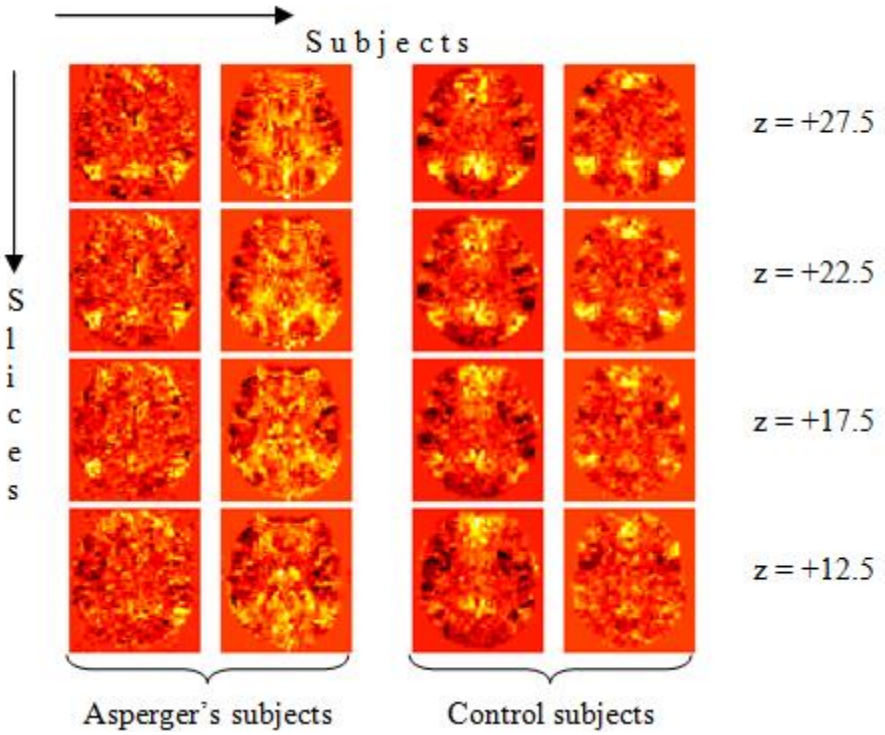

Figure 4.2: Correlation maps of four slices (z=+10 to z=+30) of four subjects (2 subjects with Asperger's disorder and 2 control subjects) representative of functional connectivity

A spatial mask was generated for each slice location to include only common brain regions among all subjects. This was accomplished in MATLAB by performing logical AND between same slice of every subject and then performing dilation to form contiguous maps. In addition to the masking technique, we explored three additional preprocessing steps and their combinations in order to improve the prediction accuracy of the SVM. The details of each of these preprocessing steps and the motivation behind using them are discussed below.

a) Absolute z-score values

Functional connectivity is a measure of the synchrony of activation or deactivation between different neural assemblies. Irrespective of whether the timecourses of voxels are positively or negatively correlated, connectivity can be viewed as a measure of the strength (magnitude) of this association. In this respect, the absolute z-score value could prove to be a better feature than

the original z-score value to represent the strength of the DMN connectivity. Especially in our problem, it was reported in [17] that the DMN connectivity was weaker in Asperger's Disorder subjects as compared to healthy control subjects. Hence, we deduced that training SVM with absolute correlation values may help improve the classification accuracy. In this preprocessing step, absolute values of the correlation maps were taken in order to represent the connectivity between a particular brain voxel and the PCC.

b) Spatial smoothing of correlation maps

In spite of normalization, there might be variability in the exact shape and size of brain structures in individual subjects. For example, edges were observed in the correlation maps, suggestive of a sharp transition which often appeared in significantly different locations even between subjects of the same group. Such inter-subject variability may deteriorate the performance of the classifier because it may reduce the number of shared features between members of the same group. To account for this variability, spatial smoothing of each correlation map was done. Low-pass gaussian filters with various different values of FWHM were tried for this purpose and the gaussian filter with FWHM of 6 voxels gave the best classification performance and was thus finally selected.

c) Default Mode Network mask creation

The number of voxels and hence the number of correlation values used to represent each subject was ~10,000 voxels within the brain. The dimension of the data from each subject was much larger than the number of subjects, and hence it was highly likely that the problem would suffer due to "the curse of high dimensionality" [18, 19]. In other words, SVM training may overfit to the training data set, consequently degrading classification performance on test data. Therefore it would be desirable to select correlation values of only relevant brain regions to train the SVM

most efficiently. As demonstrated in [17], the default mode network activity could be used to distinguish between Asperger's Disorder subjects and control subjects. Hence an attempt was made to use a spatial mask covering only the default mode network. The default mode network was identified from the correlation maps of the 17 control subjects by including only those voxels that had a significant correlation value in at least 3 subjects.

4.4.2.5 SVM algorithm

SVM training and testing was implemented using LIBSVM [29] in MATLAB. In terms of implementation a linear kernel was used (Different kernels including the sigmoid and radial basis function were tried in conjunction with various combinations of SVM training parameters, but none succeeded in improving the prediction accuracy (data not shown)).

Because the number of subjects was small (7 Asperger's, 17 controls), leave-one-out cross validation (LOOCV) was used to estimate the accuracy of the classification process. Each time the SVM was trained on the maps of 23 subjects and then tested on the remaining subject's map. Out of the 24 cases, the total number of times the test subject's class was correctly predicted was recorded.

The preprocessing steps mentioned above were applied in all possible combinations and the SVM algorithm was then used to train and test this data using LOOCV as described above. SVM training involves generation of discrimination maps or weight vector maps in which different weights are assigned to each voxel. These weights can be interpreted as the strength of contribution of the respective voxel to the classification. Thus, in our study, weight vector maps are indicative of parts of the brain that are most dissimilar between the two groups and hence most discriminative.

4.4.3 Results

Table 4.1 summarizes the prediction accuracies of the cross validation with different combinations of preprocessing steps. As observed from the table, the highest prediction accuracy was observed when a combination of all three preprocessing steps was used. Application of one or two of these steps sometimes degraded the prediction accuracy. For example, when the maps were only spatially smoothed, the classification accuracy reduced from 79% to 66% and when these maps were spatially smoothed and also a default mode network mask was used, the performance deteriorated from 79% to 75%. These two cases have in common that absolute correlation values are not used, and this reaffirms our conjecture that the classification should be based on the strength of DMN connectivity.

	Preprocessing Combinations							
Preprocessing Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Absolute correlation		✔			✔	✔		✔
Spatial smoothing			✔		✔		✔	✔
DMN mask				✔		✔	✔	✔
Prediction Accuracy	**79 %**	**83 %**	**66 %**	**79 %**	**92 %**	**83 %**	**75 %**	**96 %**

Table 4.1: Prediction accuracy results for different combinations of preprocessing methods. (DMN – Default mode network)

Figure 4.3 depicts a weight vector map obtained after training the SVM using absolute correlation values and spatial smoothing on z-maps of all 24 subjects (with no default mode network masking). Only the top 10% of the weights are displayed; these weights are contained in the default mode network. This reaffirms the claim that brain regions belonging to the default mode network are vital in classifying Asperger's Disorder subjects from control subjects.

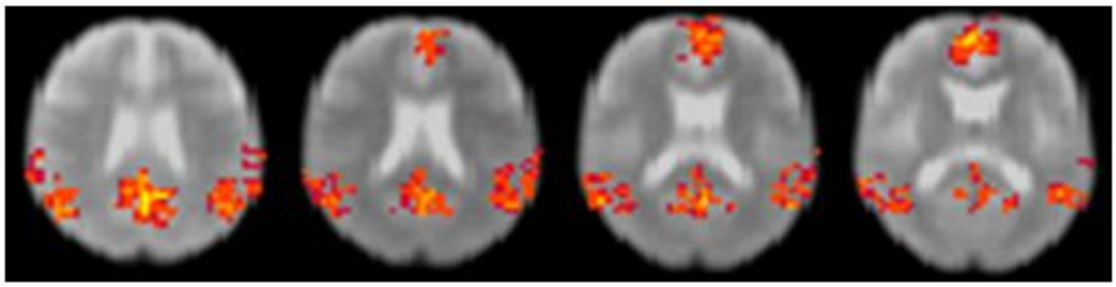

Figure 4.3: Weight vector maps for four slices (z=+12.5, z=+17.5, z=+22.5 and z=+27.5), using absolute correlation and spatial smoothing (Combination 5). Only the top 10% of positive weights have been displayed.

4.4.4 Conclusions

The present study introduces a novel framework which involves the use of support vector machines as a machine learning tool implemented to classify Asperger's Disorder subjects from healthy control subjects. Support vector machine was trained with correlation values between the resting state fMRI time series of each voxel and that of a seed voxel taken from posterior cingulate cortex. In the leave-one-out cross validation on a data set of 7 Asperger's Disorder subjects and 17 healthy control subjects, classification was successfully performed with high prediction accuracy up to 96%, which was achieved with the application of preprocessing steps including absolute correlation values, spatial smoothing of correlation maps, and use of a default mode network mask. A previous study [17] reported that relative to controls, the disordered population showed an alteration in functional connectivity. Specifically, the disordered group showed weaker connectivity between the posterior cingulate cortex and medial prefrontal cortex. Our findings are consistent with these results as depicted in the classification accuracy numbers and the weight vector maps.

Future work would include verifying these findings on a much larger dataset. Also, a more automated and refined process of feature selection such as the automated segmentation of brain regions suggested in [30] would be helpful.

4.5 Concluding remarks

In this chapter, we looked at the use of support vector classification algorithm for enabling the categorization of subjects into healthy controls or disordered patients. Resting state functional connectivity maps were used with posterior cingulate cortex chosen as the seed for connectivity analysis. This study can be further extended by including the use of support vector regression to relate the clinical severity of autism to the fMRI resting state functional connectivity measure. In that investigation, a subject with Asperger's Disorder would be labeled with a behavioral score assessed by clinicians which indicates the clinical severity of the disorder. After training on these data and labels, the model built by the SVM would predict a score for the test data corresponding to the clinical severity of their disorder. Quantification of the severity of Asperger's disorder from fMRI data by using support vector regression could prove to be a valuable clinical tool to assist physicians in their diagnoses. This may also help elucidate the neural cause of the disorder in further detail.

REFERENCES:

[1] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. Frackowiak, "Functional connectivity: the principal-component analysis of large (PET) data sets.," *J. Cereb. Blood Flow Metab.*, vol. 13, no. 1, pp. 5–14, Jan. 1993.

[2] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional Connectivity in the Motor Cortex of Resting," MRM, no. 34, pp. 537–541, 1995

[3] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 27, pp. 9673–8, Jul. 2005.

[4] V. G. van de Ven, E. Formisano, D. Prvulovic, C. H. Roeder, and D. E. J. Linden, "Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest.," *Hum. Brain Mapp.*, vol. 22, no. 3, pp. 165–78, Jul. 2004.

[5] K. R. A Van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner, "Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization.," *J. Neurophysiol.*, vol. 103, no. 1, pp. 297–321, Jan. 2010.

[6] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 2, pp. 676–82, Jan. 2001.

[7] J. Graner, T. R. Oakes, L. M. French, and G. Riedy, "Functional MRI in the investigation of blast-related traumatic brain injury.," *Front. Neurol.*, vol. 4, no. March, p. 16, Jan. 2013.

[8] R. C. Craddock, P. E. Holtzheimer, X. P. Hu, and H. S. Mayberg, "Disease state prediction from resting state functional connectivity.," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1619–28, Dec. 2009.

[9] APA, *American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders*. 1994.

[10] K. Pelphrey, R. Adolphs, and J. P. Morris, "Neuroanatomical substrates of social cognition dysfunction in autism.," *Ment. Retard. Dev. Disabil. Res. Rev.*, vol. 10, no. 4, pp. 259–71, Jan. 2004.

[11] D. E. Welchew, C. Ashwin, K. Berkouk, R. Salvador, J. Suckling, S. Baron-Cohen, and E. Bullmore, "Functional disconnectivity of the medial temporal lobe in Asperger's syndrome.," *Biol. Psychiatry*, vol. 57, no. 9, pp. 991–8, May 2005.

[12] V. L. Cherkassky, R. K. Kana, T. A. Keller, and M. A. Just, "Functional connectivity in a baseline resting-state network in autism.," *Neuroreport*, vol. 17, no. 16, pp. 1687–90, Nov. 2006.

[13] A. J. Bailey, S. Braeutigam, V. Jousmäki, and S. J. Swithenby, "Abnormal activation of face processing systems at early and intermediate latency in individuals with autism spectrum disorder: a magnetoencephalographic study.," *Eur. J. Neurosci.*, vol. 21, no. 9, pp. 2575–85, May 2005.

[14] K. M. Dalton, B. M. Nacewicz, T. Johnstone, H. S. Schaefer, M. A. Gernsbacher, H. H. Goldsmith, A. L. Alexander, and R. J. Davidson, "Gaze fixation and the neural circuitry of face processing in autism.," *Nat. Neurosci.*, vol. 8, no. 4, pp. 519–26, Apr. 2005.

[15] S.J. Weng, J. L. Wiggins, S. J. Peltier, M. Carrasco, S. Risi, C. Lord, and C. S. Monk, "Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders.," *Brain Res.*, vol. 1313, no. May, pp. 202–14, Feb. 2010.

[16] J. L. Wiggins, S. J. Peltier, S. Ashinoff, S.J. Weng, M. Carrasco, R. C. Welsh, C. Lord, and C. S. Monk, "Using a self-organizing map algorithm to detect age-related changes in functional connectivity during rest in autism spectrum disorders.," *Brain Res.*, vol. 1380, pp. 187–97, Mar. 2011.

[17] C. S. Monk, S. J. Peltier, J. L. Wiggins, S.J. Weng, M. Carrasco, S. Risi, and C. Lord, "Abnormalities of intrinsic functional connectivity in autism spectrum disorders.," *Neuroimage*, vol. 47, no. 2, pp. 764–72, Aug. 2009.

[18] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag NY, 2000.

[19] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[20] S. M. LaConte, S. Strother, V. Cherkassky, J. Anderson, and X. Hu, "Support vector machines for temporal classification of block design fMRI data.," *Neuroimage*, vol. 26, no. 2, pp. 317–29, Jun. 2005.

[21] D. Wechsler, "Wechsler Abbreviated Scale of Intelligence." 1999.

[22] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, B. L. Leventhal, P. C. Dilavore, A. Pickles, and M. Rutter, "The Autism Diagnostic Observation Schedule – Generic : A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism," *J. Autism Dev. Disord.*, vol. 30, no. 3, 2000.

[23] C. Lord, M. Rutter, and A. Le Couteur, "Autism Diagnostic Interview - Revised: A Revised Version of a Diagnostic Interview for Caregivers of Individuals with Possible Pervasive Developmental Disorders," *J. Autism Dev. Disord.*, vol. 24, no. 5, pp. 659–685, 1994.

[24] T. M. Achenbach, *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist*. 1997.

[25] Wellcome Trust Centre for Neuroimaging, "Statistical Parametric Mapping http://www.fil.ion.ucl.ac.uk/spm." .

[26] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, "Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 13, pp. 4637–42, Mar. 2004.

[27] M. J. Lowe, B. J. Mock, and J. A. Sorenson, "Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations.," *Neuroimage*, vol. 7, no. 2, pp. 119–32, Feb. 1998.

[28] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, "Mapping Functionally Related Regions of Brain with Functional Connectivity MR Imaging," Am J Neuroradiology, vol. 21, pp. 1636–1644, October 2000.

[29] C. C. Chang and C. J. Lin, "LIBSVM : A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[30] Y. Fan, D. Shen, and C. Davatzikos, "Detecting Cognitive States from fMRI images by Machine Learning and Multivariate Classification," *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2006.

CHAPTER 5

IMAGING THE CRAVING BRAIN

Craving is defined as an intense desire to consume something and such powerful (typically, abnormal) desires for consumption can be found in obese individuals in the form of food craving or people who are dependent on a particular drug also have an irresistible yearning for their choice of drug. This phenomenon can have an adverse effect on the individual who may find it difficult to stay focused on their tasks and might even prove to be a hindrance in their daily routines. Therefore, there is considerable interest in understanding the psychological components of craving and studying the brain activation patterns associated with craving in the human brain.

In this section, we present findings from our studies in which functional MRI was used to study the brain of nicotine dependent subjects when they crave cigarettes. First, we explore a novel application of multi-task learning to build separate population-wide and subject-specific models of the craving brain. Then, we present the real-time fMRI neurofeedback setup used in our study to assist subjects in modulating their own brain activation patterns.

5.1 Multi-subject machine learning for brain state classification of nicotine craving

In recent times, machine learning has gained popularity in fMRI studies of brain state classification. In this setting, we typically have very few examples or data points from individual subjects to train a model. However, there is often a large amount of additional data from other subjects responding to the same stimulus that can be used beneficially. This section describes a novel multi-subject machine learning approach to brain state classification, in which we

demonstrate an improved approach to leveraging information from multiple subjects' brain activation data to build robust models for individual subjects. The approach treats each subject's brain state classification as a separate task, where each of these tasks is assumed to be related to every other task to a certain extent because all subjects are presented with the same stimulus. The algorithm then builds a classifier that learns from all the tasks simultaneously, taking advantage of the similarity between them, to improve the individual tasks. This approach is being used to investigate the brain's craving-state classification in nicotine dependent subjects. We provide promising results that demonstrate that this algorithm outperforms typical approaches, achieving classification accuracy of approximately 75%. More importantly, the proposed approach allows us to identify the subjective nature of brain activation across different subjects along with that shared among all subjects and optimally combines them to build robust classifiers of nicotine craving.

5.1.1 Introduction to the technique

Machine learning involves the use of an algorithm to facilitate learning from examples. In supervised machine learning algorithms, there is first a training phase, during which labeled input training samples are used to build a model that captures the relationship between the training samples and the corresponding labels [1]. This model is then used during the testing phase to compute an output label/prediction for any new testing data sample. Such a setup has been used with functional magnetic resonance imaging (fMRI) data to enable brain-state classification [2, 3, 4].

A typical approach to brain-state classification involves training a model for a particular subject by exclusively using data acquired by scanning that subject. However, it is often the case that multiple subjects are scanned using the identical setup and stimuli. The typical approach

fails to exploit the relatedness in brain activation of different subjects in response to similar stimuli. Alternatively, it could be advantageous to leverage the other participants' data when training a classifier for the target participant. One straightforward and common approach is to combine the target participant's training data with that of the others and train a common classifier. A major limitation of this approach is that it assumes that brain activation in response to a stimulus is uniform across all participants. In this work we propose using a multi-task approach to leverage the commonality between participants' brain activation while allowing for participant-specific activity.

The multi-task approach we employ was proposed in [5, 6] to build a classification model that takes information from several separate but related tasks simultaneously. This approach was explored to build a brain computer interface (BCI) using EEG data in [7]. In a very recent study [8], a Bayesian approach was proposed to use multi-task learning with fMRI data of subjects performing a Posner task which investigates the effect of attentional processes on episodic memory encoding. In our study, we have applied the multi-task learning framework to fMRI data of the craving population to develop a multi-subject classification scheme that helps throw light on the neural correlates of nicotine craving. In our specific case, each subject's brain state classification is treated as a separate task. Thus, applying multi-task learning in this setting yields a multi-subject machine learning algorithm designed to take advantage of the similarity in brain activation of different subjects. By exploiting this shared structure, it is possible to leverage any commonality between the subjects and build a more robust classifier. The resultant multi-subject classifier is comprised of a shared component common to all participants that identifies brain activation common to all, as well as individual participant-specific components that identify brain activation specific to each participant. This separation allows us to identify parts of the

brain that are most affected by the stimuli at the population level by means of the common component. Additionally, the participant specific components highlight the differences in how the stimuli affect the brain activity of different individuals in the population. This type of analysis provides for richer understanding and better interpretation of the effects of the stimuli on brain activity.

Though generally applicable, herein we apply the multi-subject classifier to the brain state classification of cigarette craving and non-craving brain states in nicotine dependent subjects. Smoking addiction is a major health problem, and nicotine craving can be a persistent and disturbing feature of this addiction. Studies have reported that nicotine dependence level of subjects is associated with greater blood oxygen level dependent (BOLD) fMRI activation [9, 10] and craving for cigarettes [11] in response to smoking cues. This application offers a basis for comparing the efficacy of the multi-subject approach to the standard approaches, which will serve as baselines. Additionally, it demonstrates the extra insight and interpretability afforded by the multi-subject model formulation.

The results are promising and show that it is possible to learn classifiers that predict the craving state of subjects with high accuracy. In addition to the predictive performance, an important contribution to highlight is that such methods allow us to assess the contribution of different parts of the brain to the brain-state classification. More specifically, our approach allows for the differentiation between group-level and individual participant-specific subjective activation. A permutation test performed on the multi-subject technique generates reliable maps that reveal the parts of the brain that are most vital for nicotine craving classification. Thus, our analysis presents a multivariate technique that highlights the subjective nature of an individual's brain activation in response to nicotine craving.

5.1.2 Methods

5.1.2.1 Subject Screening

Sixteen subjects with nicotine dependence participated in this study. Screening measures including expired carbon monoxide level (CO $\geq$ 12ppm), number of cigarettes smoked daily ($\geq$ 10) and Fagerström scores ($\geq$ 4) were used to ensure at least moderate smoking and level of nicotine dependence. All subjects were scanned after overnight abstinence, as verified by measuring a decrease in CO level (at least 8ppm) on the scan day compared to CO level on the screening day.

5.1.2.2 Data acquisition and paradigm

BOLD functional images were collected on a 3T GE scanner using a T2*-weighted single-shot custom spiral-in sequence. The scan parameters used are summarized as follows: TR/TE/FA/FOV=2s/30ms/90°/22cm, 64x64 matrix, 40 contiguous axial slices of 3mm thickness. Subjects were given mirrored glasses to view a rear projection screen while being scanned. The paradigm included displaying of a sequence of images that depicted smoking or non-smoking scenes. These were presented in alternating blocks to induce or suppress subject's nicotine craving. The validity of these pictorial stimuli to elicit craving has previously been demonstrated in [12]. Two such runs were collected per subject. The first run was used for training the classification model, whereas the second run was used to assess the effectiveness of the model. Details of the paradigm can be summarized as follows: (20s blocks, each block with 5 pictures for 4s each; 16 repeats of alternating blocks of craving and non-craving descriptors, with 4s static fixation image in between each block; 384s total time).

5.1.2.3 Preprocessing steps

A custom MATLAB code was used for k-space outlier removal and spiral reconstruction. SPM8 [13] was used to perform slice timing correction, which corrects for differences in acquisition time between slices during sequential imaging. It performs a phase shift, resulting in each time series having the values that would have been obtained had the slice been acquired at the beginning of each TR. SPM motion correction was performed to reduce the effects of head motion that might have corrupted the data. To allow for spatial correspondence between brains in this multi-subject analysis, the brains were warped to an MNI template space. Spatial smoothing was then done using a Gaussian smoothing kernel with FWHM of 8 mm to mitigate the effects of inter-subject variability. Furthermore, the time course of each voxel was normalized by subtracting its mean over time and dividing by its standard deviation.

5.1.2.4 Features and examples

A classifier is a function that takes a set of feature values representing an example as input and predicts the class (label) that the particular example belongs to as an output. More specifically, if x is an example with features $[x_1, x_2, x_3, \ldots]$ and the class label is denoted by $y=(\pm 1)$, then the classifier is a function f( ) that computes the label for a given input, i.e. $y=f(x)$. In our study, at each time point, a BOLD brain activation volume is acquired. Each such volume is used as a separate example in which the voxel grey scale intensities act as features. Depending on whether the subject was looking at a smoking or non-smoking picture stimulus, a label of +1 or -1 is associated with each example.

Data acquired during the first run are used as training data. In the training phase, a mapping is learned from the training examples to the respective class labels and a classifier is built. In the testing phase, this model is used to predict the class of a previously unseen example

from the testing data acquired during the second run. Classifier performance is calculated as the ratio of the number of correctly classified test examples to the total number of test examples.

5.1.2.5 Dataset dimensionality

Each of the acquired 3-dimensional volumes, which act as training examples, was of size [53x63x46] voxels. Initially, this accounted to 153,594 features that were then reduced to 54,438 features by excluding all voxels that fell outside the brain region. Disregarding the brain images captured when the subjects were looking at the fixation cross that was displayed between blocks, we obtained a total of 160 examples per run with two such runs per subject.

5.1.2.6 SVM classification

To create a temporal brain-state classifier across multiple subjects we compare three commonly used SVM approaches, which will serve as touchstones to the multi-subject SVM approach discussed herein. In all of the subsequent cases described, each subject's first run is used for the training phase, whereas, their second run is used for testing. The first touchstone, T1E1 (train on 1, test on 1), treats each participant independently and trains a classifier only on their brain activity, disregarding any commonality between subjects that arises from the shared paradigm. The second, T16E1 (train on all 16, test on 1), combines all participants' activity and trains one classifier that is generally applicable to all participants assuming that brain activity is exactly similar across all participants. The third, T15E1 (train on other 15, test on left out 1), is similar to T16E1 and trains one classifier using all the participants' data *excluding* the target participant's training data. The classifier is then tested on the target participant's second run. This final scenario simulates a situation where a classifier trained on previously scanned subjects is used to classify the brain state of a new subject.

These three baseline approaches are compared to the multi-task approach, proposed in [5, 6]. This technique, when applied such that each subject's data defines a task, can be manipulated to address the spectrum between the T1E1 and T16E1 (i.e. entire range from least dependence on other subjects' data to most dependence). The resultant multi-subject (MS) classifier combines, in a principled manner, a participant-specific classifier with a generally applicable one. With regard to the smoking- vs. control-stimulus task, the MS classifier outperforms the baselines by training a classifier that emphasizes participant-specific brain activity while leveraging the commonality between brain activities of other subjects responding to the same paradigm.

As a quick refresher, in traditional two-class SVM classification approach, a separating boundary between the two classes of examples (e.g. +1 and −1) is learned such that the margin between the data points and boundary is maximal. In higher dimensions, this boundary manifests itself as a hyperplane. This separating hyperplane generated by the SVM algorithm is orthogonal to the weight vector **w.** The SVM then uses the sign of the decision function $f(x)=\mathbf{w}^T\mathbf{x}$ to classify any data point x represented by the feature vector **x** into one class or the other. Thus, the classifier is parameterized by **w,** which can be solved for by using the following optimization:

$$\begin{aligned} &\min_{\mathbf{w}} \| \mathbf{w} \|^2 + C \sum_{i=1}^{n} \xi_i \\ &s.t.\ \forall i \in \{1, 2, ..., n\}: \\ &\mathbf{y}_i \mathbf{w}^T \mathbf{x}_i \geq 1 - \xi_i \ \& \ \xi_i \geq 0 \end{aligned} \qquad (1)$$

where **w** is the normal vector to the hyperplane, $\mathbf{y}_i$ are the known input class labels, $\mathbf{x}_i$ are the input feature vectors, C is the trade-off parameter used to penalize misclassifications and ξ_i are the non-negative slack variables which measure the degree of misclassification of the input data $\mathbf{x}_i$.

In this study, the MS classifier used is the multi-task SVM classifier proposed in [5, 6] where each task *t* represents an individual subject's brain activation in response to a common craving paradigm. This multi-task SVM classifier learns, in a coupled manner, solutions for T tasks using a separate classification function for each task *t*, $f_t(x) = \mathbf{w}_t^T\mathbf{x}$. The learnt classifier $\mathbf{w}_t$ for each task *t* is defined as:

$$\mathbf{w}_t = \mathbf{w}_0 + \mathbf{v}_t \qquad (2)$$

where $\mathbf{w_0}$ is shared across all tasks and $\mathbf{v}_t$ is specific to each task *t*. When the vectors $\mathbf{v}_t$ are large relative to $\mathbf{w_0}$, the task-specific components dominate the shared component and each task may have a very different classifier. When $\mathbf{v}_t$ is small relative to $\mathbf{w_0}$, all of the tasks have very similar classifiers.

To obtain these hyperplanes, one can solve the optimization problem:

$$\begin{aligned} &\min_{\mathbf{w}_0,\mathbf{v}_t,\xi_{it}} \; C\sum_{t=1}^{T}\sum_{i=1}^{m}\xi_{it} + \frac{1}{\lambda T}\sum_{t=1}^{T}\| \mathbf{v}_t \|^2 + \frac{1}{1-\lambda}\| \mathbf{w}_0 \|^2 \\ &s.t.\ \forall i \in \{1, 2, ..., n\} \;\&\; t \in \{1, 2, ..., T\}: \\ &\mathbf{y}_{it}(\mathbf{w}_0 + \mathbf{v}_t)^T \mathbf{x}_{it} \geq 1 - \xi_{it} \;\&\; \xi_{it} \geq 0 \end{aligned} \qquad (3)$$

where *T* is the number of subjects (tasks) used for training and m is the number of samples or training volumes per task. $\mathbf{y}_{it}$ are the known input class labels, $\mathbf{x}_{it}$ are the input feature vectors for subject t. C is the trade-off parameter used to penalize misclassifications and ξ_i are the non-negative slack variables which measure the degree of misclassification of the input data $\mathbf{x}_{it}$. The regularization parameter $\boldsymbol{\lambda}$ determines the cost parameter of the SVM as well as enforces the relatedness of the tasks. Thus, it models relations among the tasks by acting as a task coupling parameter and quantifies how related the tasks are.

LIBSVM [14], a Library for Support Vector Machines, was used to perform the SVM classification with the pre-computed multi-task kernel.

The multi-task kernel is defined as:

$$K\{(x,t)(s,q)\} = (1 - \lambda + \lambda T \delta_{tq}) x^T s \qquad (4)$$

where data point x from task *t* is represented as an ordered pair (x,*t*), and data point s from task *q* as the ordered pair (s,*q*) and δ_{tq} is 1 when $t=q$ and 0 otherwise. When λ equals 0, each task influences every other task heavily and makes the problem similar to building one common classifier for all tasks together (T16E1) whereas when λ equals 1, it implies that the tasks are all learned independently and that an individual SVM classifier is learned per task (T1E1). Thus, the choice of λ is an important decision in multi-task problems.

5.1.2.7 Experimental iterations

To determine how our classification scheme performed with a varied number of examples from subjects, we designed a specific setup, referred to as Setup 1, for the analysis in which four cases were examined using a reduced number of samples. In the first case, only the first 25% (i.e. first 40 out of the total 160) examples from all the subjects' data were used to train a model. In the subsequent cases, 50%, 75% and 100% (80, 120 and all 160 examples respectively) of all subjects' data were used.

It was noted, though, that in an experimental setting, apart from the new subject being scanned (target subject), all the training examples (entire 100%) from subjects other than our target subject would already be available. This data could, therefore, be leveraged to improve our classifier. In this regard, Setup 2 was designed to study the effect of varying the number of examples from the target subject used to train a model while using all existing data from the others. Similar to Setup 1, the analysis was performed in four cases. In the first case, only the first 25% (i.e. first 40 out of the total 160) examples from the target subject's data were used along with the entire 100% of all the other subjects' data (all 160 examples) to train a model. In

the subsequent cases, 50%, 75% and 100% (80, 120 and all 160 examples respectively) of the data from the target subject were used along with the entire 100% of all the other subjects' data (160 examples).

These experimental setups were designed to examine whether a fewer number of examples suffice to learn robust classifiers that achieve classification accuracy comparable to that of classifiers trained by using all the examples available. Additionally, to simulate a realistic deployment scenario, when selecting a subset of the data in the iterations we used the first portion of examples instead of picking them at random. This approach to subset selection might also account for effects such as fatigue and habituation.

5.1.2.8 Parameter selection

In the SVM formulation, there is a tradeoff parameter C which allows the user to adjust the penalty imposed on each misclassification as seen in (1) and (3). This, in turn, governs the number of training misclassifications permitted. If C is set too high there is a huge penalty on misclassifications, which generates a model that is more rigid and less generalizable. On the other hand, if it is set too low, there is not much penalty on misclassifications which may generate an inept model. Hence, as is typically the case, C is chosen based on a validation set. Values of $C=2^{\wedge [-20:1:20]}$ were considered in both the standard SVM and the MS-SVM. For the multi-task kernel, the coupling parameter λ was also chosen using a validation set, and was allowed to vary over the entire range from 0 to 1, with smaller steps of 0.025 from 0 to 0.4 and larger steps of 0.1 from 0.4 to 1. A leave-one-subject-out cross validation setup was used to choose the best C and λ parameters for each participant.

5.1.2.9 Permutation tests

The SVM algorithm generates a separating hyperplane which is orthogonal to the weight vector **w**. This weight vector defines the direction in which samples of the two classes differ most from one another. Thus, it is representative of the most discriminatory regions of the brain. This vector when mapped back into original image space generates the discriminating volume also called the weight vector map. Each voxel is assigned a separate weight and since SVM is a multivariate pattern analysis technique, all weight values depend on all other weight values. The weight vector map is thus a representation of the voxels that are most vital to the classification. The magnitude of the absolute value of each voxel weight determines its importance in discriminating the brain states and the most important voxels for discrimination of internal emotional/cognitive states can be inspected by merely thresholding the obtained weight vector map. However, to assess the reproducibility of these spatial patterns, a permutation test has been employed as discussed later.

Permutation tests are nonparametric techniques that empirically estimate the distribution of a statistic under a null hypothesis and have been used with fMRI data previously in [15, 16]. The null hypothesis suggests there are no differences between the two brain states and thus the labels assigned to each example, delineating them as belonging to one class or another, are inconsequential. The alternate hypothesis proposes that the assigned class labels, in fact, are related to the classification and much better than random. One can estimate the distribution of weights assigned to each voxel under the null hypothesis by randomly permuting the class labels 2000 times and then training the SVM each time with a different permutation of labels. In each instance, the weights are normalized to have unit standard deviation. The SVM training is also done once with the known correct non-permuted labels. Now, for each voxel, the p value under null hypothesis can be calculated as the ratio of number of times (out of 2000) that the voxel

weight assigned to it is greater than or equal to the weight assigned to it when training with original non-permuted labels. If this number is smaller than 100, then that voxel is likely to be predictive of the class label with a significance level of 5%. The weight vector maps shown in the results section display all significant voxels with p value < 0.05.

5.1.3 Results

5.1.3.1 Classification accuracy

Figure 5.1 shows a plot of the mean classification accuracy across all 16 subjects against the percentage of examples used for training the classifier model using Setup 1. As expected, T1E1 (blue) performs better with an increase in the number of examples used for training the model. Since T15E1 (green) uses data from every subject except the target subject to build the model, it does not perform as well. T16E1 (red) performs better with an increase in the number of examples but MS (black) outperforms all the other approaches.

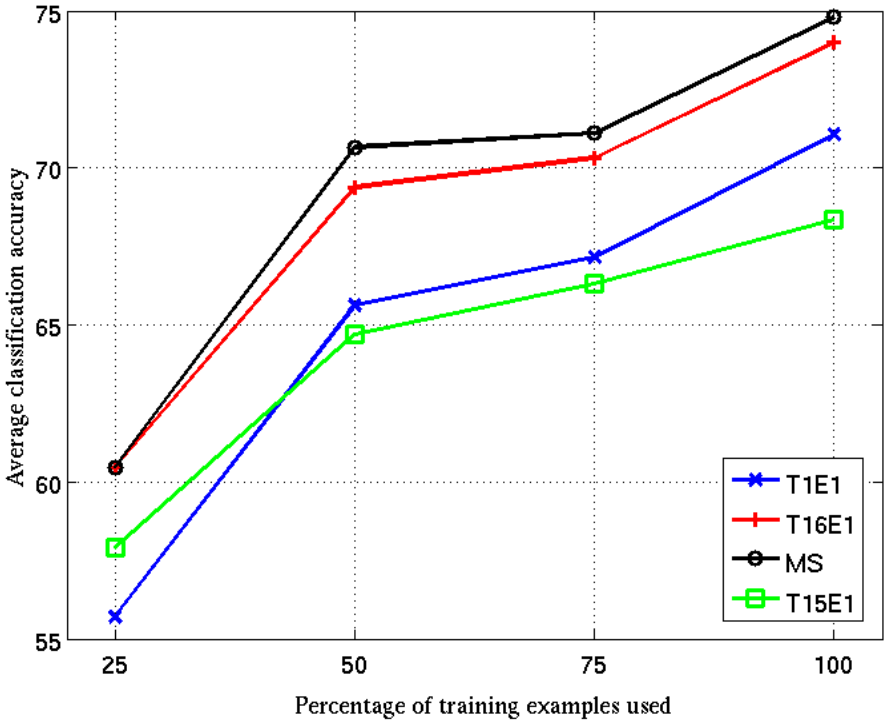

Figure 5.1: Setup 1 plot of mean classification accuracy across all 16 subjects against percentage of training examples used to train the model for each of the classification approaches.

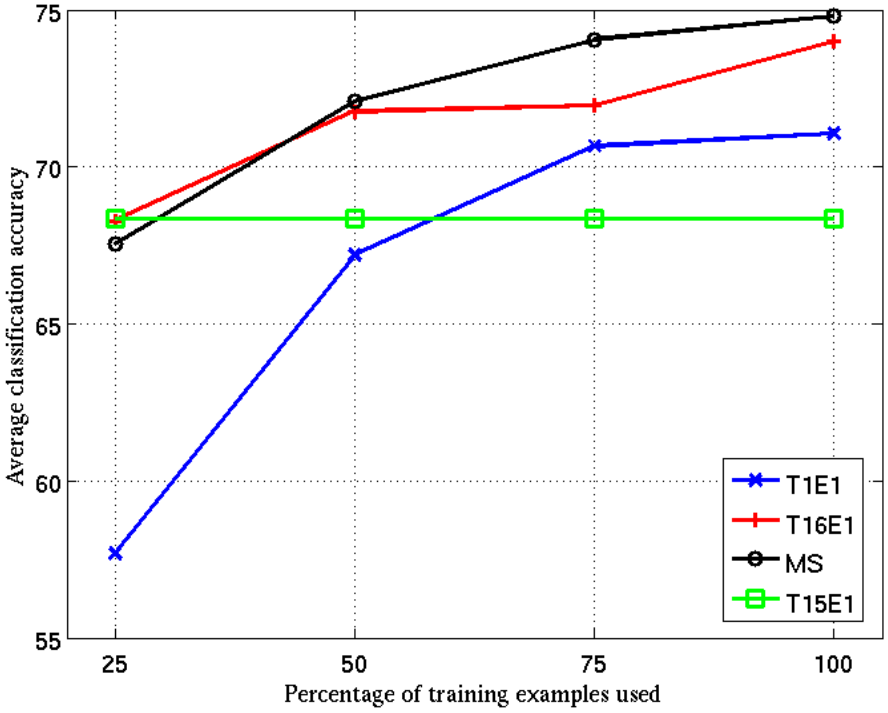

Figure 5.2: Setup 2 plot of mean classification accuracy across all 16 subjects against percentage of training examples used to train the model for each of the classification approaches.

Figure 5.2 shows a plot of the mean classification accuracy across all 16 subjects against the percentage of examples used for training the classifier model using Setup 2. It is clear from the plot that including more training data from the target participant improves classifier performance. Also, including sufficient training data (in this case, $\geq$ 50% or 80 examples) from the target participant is important for good performance. T15E1 (green), which has no training data from the target participant, performs worst in most cases.

More importantly, it can be observed that T16E1 (red) performs better than MS (black) when we have less information from the target subject while training a model. But when we include more data (50%, 75% and 100%, i.e. 80, 120 and 160 examples), then the multi-subject approach performs best. Additionally, we observe that MS with 75% of the participant's data approximately matches the performance of MS at 100% and outperforms T16E1 with 100%, meaning that by using the MS approach we can acquire fewer training data examples from the target participant with comparable results.

The classification accuracies obtained for each individual subject by using each of the four methods (T1E1, T15E1, T16E1 and MS) are summarized in Table 5.1. These results exhibit that for some subjects, training a model using only their data (T1E1) gives better classification accuracy. This is perhaps an indication that these subjects' activations are significantly different from the general population and thus introducing other participants' data worsens their model. For few others, training a model on all the subjects' data pooled together (T16E1) helps train a better model pointing towards the likelihood that these subjects' brain activation is very similar to that of the general population. But for a majority of the subjects, leveraging the classifier by using the other subjects' data along with the target subject's training run (MS) helps build the best classifier.

Subject #	T1E1	T15E1	T16E1	MS
1	56.88	61.88	*68.75*	66.25
2	72.50	71.88	75.63	*79.38*
3	69.38	71.25	*71.25*	70.63
4	74.38	76.88	78.75	*80.00*
5	81.25	68.13	80.00	*83.13*
6	76.88	68.13	*77.50*	76.88
7	66.88	68.75	71.25	*74.38*
8	85.00	73.13	82.50	*85.63*
9	*73.13*	65.00	65.00	67.50
10	58.75	61.25	*66.25*	64.38
11	70.00	70.63	78.75	*79.38*
12	78.75	70.63	79.38	*83.13*
13	70.63	65.63	70.00	*72.50*
14	*77.50*	72.50	73.13	71.88
15	64.38	74.38	*81.88*	79.38
16	60.62	53.75	*63.75*	62.50
Mean	71.06	68.36	73.99	74.81

Table 5.1: Comparison of classification accuracies for each individual subject across each of the four methods. The maximum accuracy in each row is italicized for emphasis.

5.1.3.2 Weight vector maps

Beyond improved classification, the proposed MS method enables us to generate discriminatory weight maps, described in the methods, to examine the population-wide as well as subjective effects of nicotine craving on brain activation. All the displayed maps are obtained by permutation tests and only voxels that are predictive with a significance level of $p<0.05$ have been superimposed on the MNI brain template. The population-wide common component $\mathbf{w}_0$ is represented in red while the subject-specific component $\mathbf{v}_t$ is shown in yellow and their intersection is displayed in green. Figure 5.3-(a) corresponds to the mean $\mathbf{w}_0$ image, whereas, Figure 5.3-(b) corresponds to $\mathbf{w}_0 + \mathbf{v}_t$. A representative subset of subjects and slices are selected for ease of visual representation. These clusters were then compared to the Automated Anatomical Labeling (AAL) template defined in [17] to find respective anatomical regions.

Names of all the anatomical regions that show clusters in the maps have been summarized in Table 5.2.

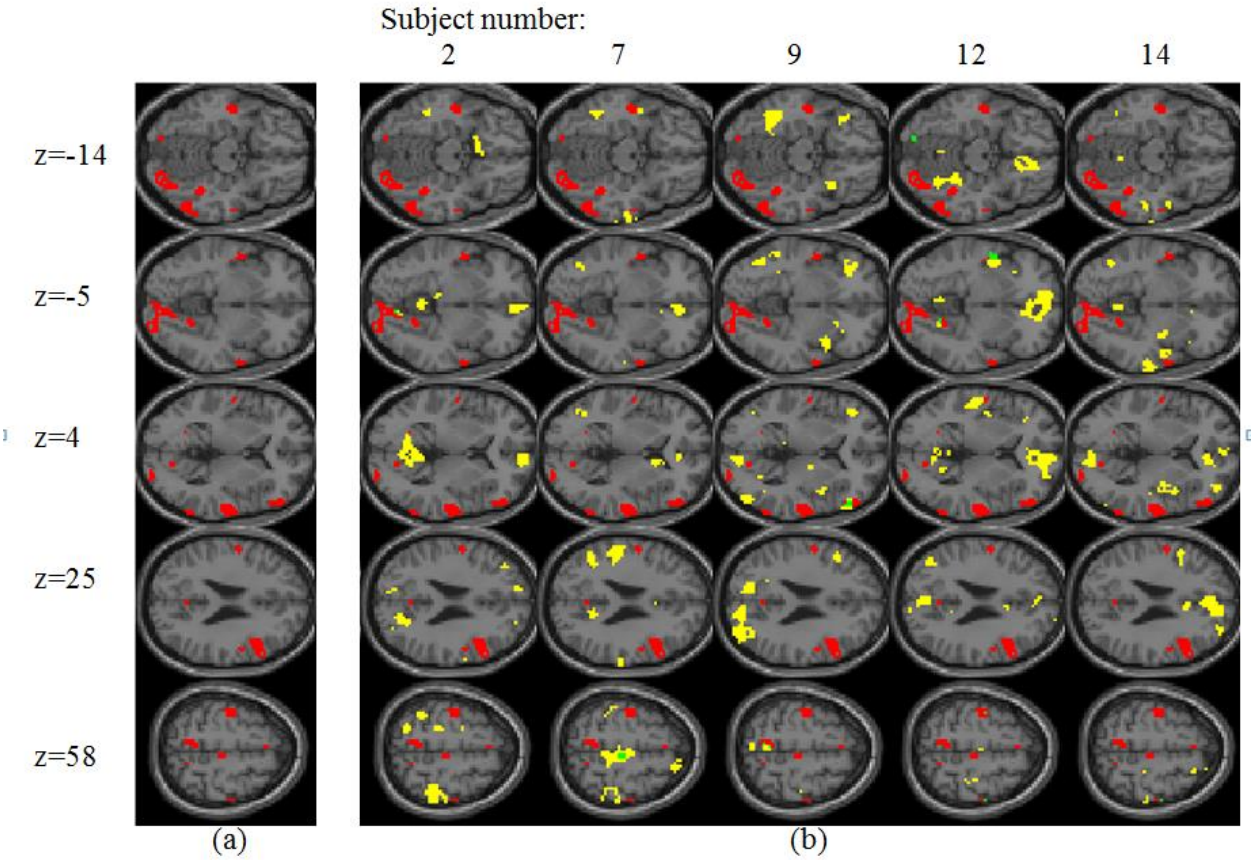

Figure 5.3: Example of weight vector map obtained using permutation tests ($p<0.05$) on multi-subject SVM superimposed on MNI brain template. Population-wide component w_0 [red] and individual participant-specific component v_t [yellow] and overlap [green] (a) w_0 (b) $w_0 + v_t$ Each row corresponds to a different slice (z=-14, z=-5, z=4, z=25, z=58, from top to bottom) of the MNI template and each column corresponds to a different subject (w_0+v_2, w_0+v_7, w_0+v_9, w_0+v_{12}, w_0+v_{14}, from left to right)

The $\mathbf{w}_0$ image in Fig. 3-(a) shows lingual gyrus activity, along with cuneus and precuneus activation near the occipital cortex. Precentral and postcentral sulcus activation is also evident along with supplementary motor and insular activity. Each column of Fig. 3-(b) corresponds to a single participant ($\mathbf{w}_0+\mathbf{v}_t$) and shows a different pattern of activation. Various degrees of frontal and especially orbitofrontal (OFC) activation are observed in the maps for subjects 2, 9 and 12.

Larger clusters in the anterior cingulate cortex (ACC) are seen in the map belonging to subjects 12 and 14 whereas a smaller cluster is observed for subject 7. Different proportion of pre- and post-central sulcus activation is seen in almost all subjects' maps. Subjects 7 and 9 show clusters in the precuneus region whereas subjects 2 and 12 show similar clusters in the cuneus region and also the lingual gyrus. In addition, subjects 9, 12 and 14 also show insular activation. The meaning of each of these observations has been further explored in the DISCUSSION section.

	w_0	v_2	v_7	v_9	v_{12}	v_{14}
z = -14	Lingual gyrus			Orbitofrontal cortex, insula	Orbitofrontal cortex	
z = -5	Lingual gyrus, insula, occipital cortex	Orbitofrontal cortex	ACC	Orbitofrontal cortex, insula	Orbitofrontal cortex, insula, lingual gyrus	
z = 4	Cuneus	Lingual gyrus	Frontal cortex, caudate	Occipital cortex	Lingual gyrus, frontal cortex, ACC	Insula, ACC, frontal cortex
z = 25	Precuneus, pre and post central sulcus	Cuneus	Precuneus,	Cuneus, occipital cortex	Cuneus, occipital cortex, ACC	ACC, frontal cortex
z = 58	Pre and post central sulcus, supplementary motor area	Pre and post central sulcus	Precuneus, pre and post central sulcus	Precuneus	Precentral sulcus, supplementary motor area	Postcentral sulcus

Table 5.2: Summary of different brain regions that show clusters in the weight vector maps for different subjects. Each row corresponds to a different slice as shown by the corresponding z-value of the MNI template. First column corresponds to the population-wide component of the weight vector map whereas all the other columns belong to individual subjects.

Figure 5.4 provides a visual representation summarizing the statistically significant clusters of voxels in the weight maps of all subjects belonging to all different regions of interest (ROI). The intensity of each block signifies the normalized cluster size belonging to a particular region as denoted by the AAL template. Each column represents a different ROI and for each region, the cluster size has been normalized by the standard deviation across all subjects. The top most row corresponds to the population wide component, whereas, every subsequent row corresponds to an individual subject. Most of the subjects show larger clusters in the frontal

region with subject 13 (v_{13}) showing the largest cluster. Contrarily, most subjects have smaller clusters in the ACC except subjects 6, 12 and 14 ($v_{6,}$ v_{12} and v_{14}) have large clusters.

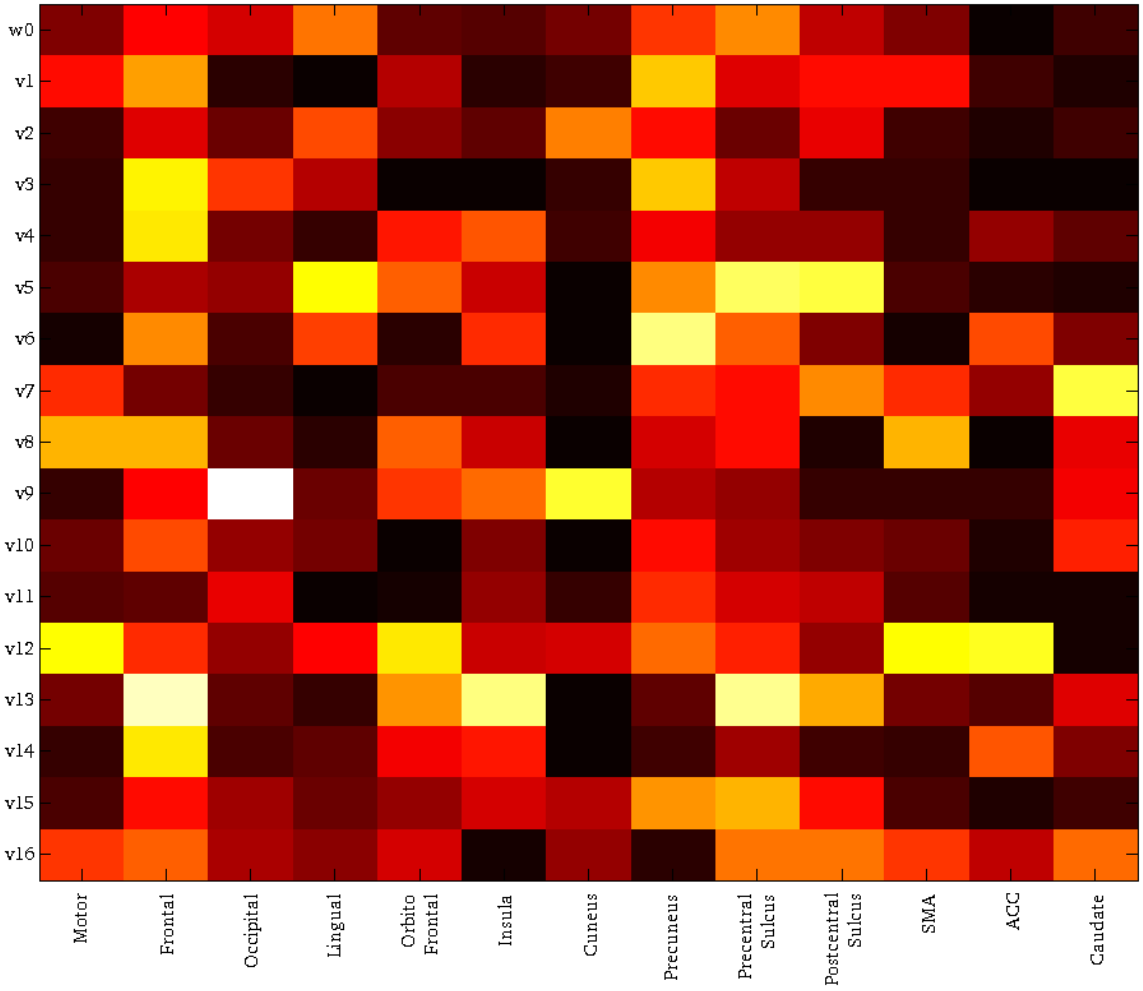

Figure 5.4: Visual representation of the summary of cluster sizes of statistically significant weights ($p<0.05$) found in different regions of interest (denoted along each column) for different subjects (denoted along each row). Each column corresponds to a different ROI as denoted (in accordance to the AAL template). First row corresponds to the population wide component of the weight vector map (w_0), whereas, all the other rows belong to individual subjects (v_t). Intensities for each ROI are normalized by the standard deviation across all subjects. Most of the subjects show larger clusters in the precuneus region with subject 6 (v_6) showing the largest cluster whereas subjects 13, 14 and 16 have small clusters.

5.1.4 Discussion

The population-wide significant weight vector map shows clusters in various regions of interest that have previously been cited in the craving literature. Lingual gyrus activity has been associated to vision processing and especially for encoding visual memories [18]. Previous functional studies have linked precuneus activation to self-consciousness and reflective self-awareness as well as processing of episodic memories [19]. The insula has been associated with explicit motivation to take drugs, i.e. urge to smoke [20]. The supplementary motor area (SMA) is a part of the brain that contributes to the control of movement. The precentral sulcus and postcentral sulcus lie on either side of the central sulcus in proximity to the motor cortex and somatosensory cortex. These regions are associated with the act of motor imagery, which is a mental execution of a movement without any explicit movement [21]. Involvement of these regions may be linked to the subjects' imagery of their hand action while smoking a cigarette.

The individual, subject-specific weight vector maps displayed significant clusters in varied regions that have been associated with cigarette craving. A recent study established that cue-induced craving signal correlated most with activity in the medial orbitofrontal cortex (mOFC), which encodes the subjective value of the drug based on content of the cue [22]. Other brain regions such as the ACC, insula and supplementary motor areas have been shown to be activated in response to the presentation of drug cues especially during abstinence [23].

In this work, the results demonstrate that using the MS approach, one can train brain-state classifiers that achieve 75% average performance on nicotine craving classification on new participants. Additionally, the MS approach allows for training better models while using less participant-specific data. These classifiers can thus be quickly deployed on new participants because less training data need to be collected, saving on the expensive and time-consuming

process of collecting a large amount of participant-specific training data. This adds to patient comfort and helps reduce the undesirable effects of patient motion caused due to long scan times. The present method not only has predictive capability, but also affords greater interpretability regarding brain function. The MS approach provides a multivariate pattern analysis technique that utilizes the population-wide brain activation component and leverages the individuals' subject-specific component to build more robust classifiers. The subjectivity of brain activation is therefore an important confounder that should not be overlooked when performing brain state classification. The otherwise individually heterogeneous regions identified in the individual $\mathbf{v}_t$ maps in the RESULTS have been previously associated with drug cue-induced craving responses in the literature [24, 25, 26, 27], further corroborating the validity of our approach. The individually-variable nature of brain activation warrants further study, however one possible explanation could be that subjects had significantly different perspectives and motivations towards smoking cigarettes and nicotine craving, or various levels of addiction.

As demonstrated, our proposed multi-subject learning approach provides a novel means to train brain-state classifiers that perform well on brain volumes from new participants and sheds light on neural correlates of smoking cue-induced craving. Moreover, by exploiting the relatedness of brain-activation patterns between the participants, these generalizable classifiers can be trained with a small amount of training participants and samples. This approach can be used to predict the craving state of a nicotine dependent subject by examining their fMRI brain activation data. These findings also encourage the possibility of using a neurofeedback mechanism to help subjects self-regulate their craving by adopting a method similar to that described in [28], but with greater capacity to incorporate individual differences in cue-induced brain regional activation.

This MS approach, whose utility we presented using nicotine craving, is generally applicable to other brain-state classification tasks where data are collected from multiple participants presented with the same stimuli, which is typical of most fMRI studies. By leveraging the existing information available from other subjects' scans, this approach helps build robust classifiers, which may improve our understanding of the neural substrates of various brain functions.

5.2 Real-time neurofeedback

Technological advancements have made it possible to increase the speed and computing power of machines tremendously. This has allowed reconstruction algorithms to be extremely fast and led to the possibility of a real-time fMRI setup. In Chapter 2, we mentioned a number of studies reporting the possible applications and results of using real-time fMRI neurofeedback to assist subjects in controlling their own brain activations. In this section, we have presented our setup for real-time fMRI and its application for modulating craving in nicotine dependent subjects.

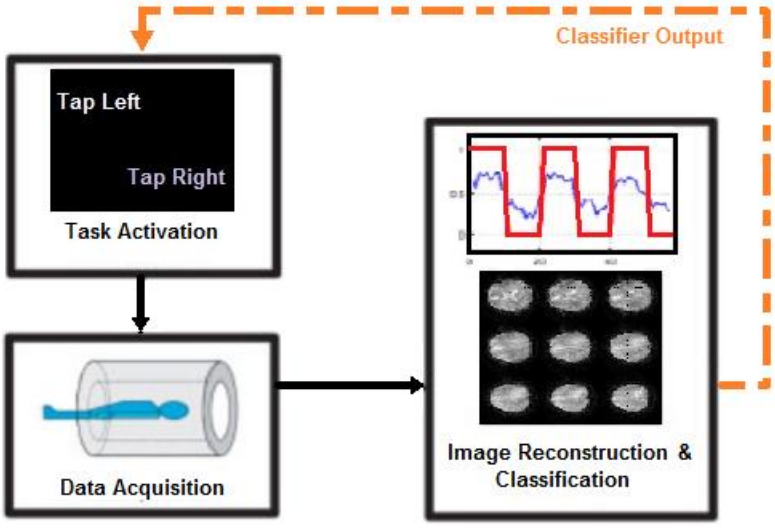

Figure 5.5: Illustration of real-time neurofeedback experiment as in [4]

Fig. 5.5 shows a simplified illustration of the real-time neurofeedback experiment that has been developed in this study. While the subject is lying inside the scanner, they use mirrored glasses to view a rear projection screen. According to the experimental design, the screen shows different cues and the subject is expected to respond to these cues. The subject's fMRI images are acquired and reconstructed after every time of repetition (TR, typically = 2s).

After the training run, all these brain volumes are used as labeled examples to train a model of the subject's brain. During the testing run, this model is used to classify each new brain volume reconstructed after every TR as belonging to one class or the other. This classification output can be used to modify the cues displayed to the subject. Since the subjects respond to the cues presented to them, this completes the feedback loop. Since this method involves using brain signals to modify the behavior of the subject after every single TR, it is termed as a real-time neurofeedback experiment.

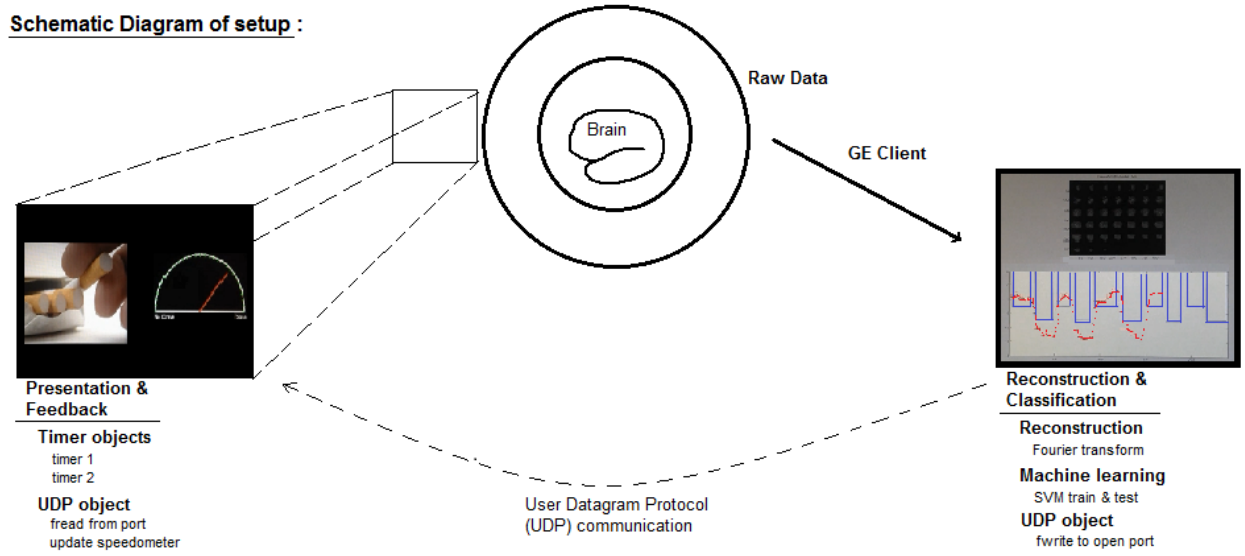

Figure 5.6: Schematic diagram of real-time setup.

Figure 5.6 shows a detailed schematic of the real-time setup which includes an MRI scanner, a presentation & feedback computer and another reconstruction and classification workstation. Whatever is displayed on the screen of the presentation computer is mirrored onto the screen behind the MRI scanner where subjects can see it using mirrored glasses.

In a standard fMRI experiment, images are captured over a five minute scan and then the series of brain volumes are reconstructed from raw data at the end. In our study, this process has been modified using a GE client to stream the raw data to the reconstruction computer after every 2 seconds. Here, a custom MATLAB reconstruction algorithm is used to reconstruct a brain volume every time raw data is received.

In our study, we use machine learning techniques to perform temporal brain state classification. The first set of acquired brain volumes are together called a training run which is used to build a model. The learnt model is then used to classify newly captured volumes in the testing phase. We use LIBSVM [14] code in MATLAB for the SVM train and SVM test.

Most groups use specialized software packages for creating the accurately timed visuals for the stimulus paradigm but we have implemented it within MATLAB by utilizing timer functions. The output of the SVM algorithm is sent from the reconstruction and classification computer via a user datagram protocol (UDP) communication channel over to the paradigm computer. The UDP package sent is either a 0 or a 1 depending on the SVM classifier output. Depending on this value, the display on the feedback computer is updated. The subject responds to this updated stimulus and thus neurofeedback is accomplished. This neurofeedback setup can be used for multiple studies that may involve a self-regulation component.

In our studies, we have used this real-time setup for controlling motor activation as well as modulation of craving related brain activation. In the case of motor activation, the subject was

instructed to tap with either left or right hand and the SVM algorithm trained a model of tapping with left hand vs tapping with right hand. This model was then used during testing phase to control the direction of motion of a speedometer needle. This finger tapping experiment was just a prototype for other more complex neurofeedback experiments like devising a cigarette craving monitor.

When used for modulating craving related brain activation, we classify between craving and non-craving temporal brain states. During the pre-scan screening session, subjects rate a series of images which may or may not have smoking descriptors. Depending on their rating of these images, the paradigm image sequence is tailored for each subject. During the scan session, subjects are shown this stimulation paradigm which involves alternating blocks of smoking and non-smoking related images. Before each such block, the subjects are instructed to imagine themselves in their craving place (any physical place that they personally associate to smoking. e.g. usual hangouts, clubs, local smoke shop, etc.) or non-craving place (any physical place that dissuades them from smoking. e.g. educational institution, church, in the proximity of children, etc.) to help them modulate their craving. The subjects were instructed to crave cigarettes while looking at image cues that were smoking-related and to not crave when looking at neutral cues. SVM algorithm was used to generate a brain state classifier that distinguished craving vs non-craving brains. During the testing phase, the SVM output was used to control the speedometer needle between the two extremes of crave and no-crave.

5.3 Concluding remarks

In this chapter, we have reported findings from our studies that involved imaging the craving brain. The multi-task learning technique was applied in a novel manner to create a multi-

subject classifier which benefit from data acquired on multiple subjects. Also, we demonstrate a real-time neurofeedback setup which was used for various applications including a craving monitor which could be used to assist subjects in controlling their own cigarette craving related brain activation.

REFERENCES:

[1] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: a tutorial overview.," *Neuroimage*, vol. 45, pp. S199–209, Mar. 2009.

[2] T. M. Mitchell, R. Hutchinson, R. S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman, "Learning to Decode Cognitive States from Brain Images," *Mach. Learn.*, vol. 57, pp. 145–175, Oct. 2004.

[3] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[4] S. M. LaConte, S. J. Peltier, and X. P. Hu, "Real-time fMRI using brain-state classification.," *Hum. Brain Mapp.*, vol. 28, no. 10, pp. 1033–44, Oct. 2007.

[5] R. Caruana, "Multitask Learning," *Machine Learning*, vol. 28, pp. 41-75, 1997.

[6] T. Evgeniou and C. A. Micchelli, "Learning Multiple Tasks with Kernel Methods," *Journal of Machine Learning Research*, vol. 6, pp. 615–637, 2005.

[7] M. Alamgir, M. Grosse-Wentrup, and Y. Altun, "Multitask learning for Brain-Computer Interfaces," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 17–24, 2010.

[8] A. F. Marquand, M. Brammer, S. C. R. Williams, and O. M. Doyle, "Bayesian multi-task learning for decoding multi-subject neuroimaging data.," *Neuroimage*, vol. 92, pp. 298–311, May 2014.

[9] F. J. McClernon, F. B. Hiott, S. A. Huettel, and J. E. Rose, "Abstinence-induced changes in self-report craving correlate with event-related FMRI responses to smoking cues.," *Neuropsychopharmacology*, vol. 30, no. 10, pp. 1940–7, Oct. 2005.

[10] M. N. Smolka, M. Bühler, S. Klein, U. Zimmermann, K. Mann, A. Heinz, and D. F. Braus, "Severity of nicotine dependence modulates cue-induced brain activity in regions involved in motor preparation and imagery.," *Psychopharmacology*, vol. 184, pp. 577–88, Mar. 2006.

[11] J. M. Wertz and M. A. Sayette, "Effects of smoking opportunity on attentional bias in smokers.," *Psychol. Addict. Behav.*, vol. 15, no. 3, pp. 268–271, 2001.

[12] A. King, P. McNamara, M. Angstadt, and K. L. Phan, "Neural substrates of alcohol-induced smoking urge in heavy drinking nondaily smokers.," *Neuropsychopharmacology*, vol. 35, no. 3, pp. 692–701, Feb. 2010.

[13] Wellcome Trust Centre for Neuroimaging, "Statistical Parametric Mapping http://www.fil.ion.ucl.ac.uk/spm." .

[14] C. C. Chang and C. J. Lin, "LIBSVM : A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[15] T. E. Nichols and A. P. Holmes, "Nonparametric Permutation Tests For Functional Neuroimaging : A Primer with Examples," *Human Brain Mapping*, vol. 15, pp. 1-25, 2001.

[16] J. Mourão-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, "Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data.," *Neuroimage*, vol. 28, no. 4, pp. 980–95, Dec. 2005.

[17] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain.," *Neuroimage*, vol. 15, no. 1, pp. 273–89, Jan. 2002.

[18] W. C. Machielsen, S. A. R. . Rombouts, F. Barkhof, P. Scheltens, and M. P. Witter, "FMRI of visual encoding: reproducibility of activation.," *Hum. Brain Mapp.*, vol. 9, no. 3, pp. 156–64, Mar. 2000.

[19] H. C. Lou, B. Luber, M. Crupain, J. P. Keenan, M. Nowak, T. W. Kjaer, H. A. Sackeim, and S. H. Lisanby, "Parietal cortex and representation of the mental Self.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 17, pp. 6827–32, Apr. 2004.

[20] N. H. Naqvi, D. Rudrauf, H. Damasio, and A. Bechara, "Damage to the Insula Disrupts Addiction to Cigarette Smoking," *Science*, vol. 315, no. 5811, pp. 531–534, 2007.

[21] T. Mulder, "Motor imagery and action observation: cognitive tools for rehabilitation.," *J. Neural Transm.*, vol. 114, no. 10, pp. 1265–78, Jan. 2007.

[22] T. Hayashi, J. Hyun, A. P. Strafella, and A. Dagher, "Dorsolateral prefrontal and orbitofrontal cortex interactions during self-control of cigarette craving," *PNAS*, vol. 110, no. 11, pp. 4422-4427, 2013.

[23] D. McBride, S. P. Barrett, J. T. Kelly, A. Aw, and A. Dagher, "Effects of expectancy and abstinence on the neural response to smoking cues in cigarette smokers: an fMRI study.," *Neuropsychopharmacology*, vol. 31, no. 12, pp. 2728–38, Dec. 2006.

[24] A. L. Brody, M. A. Mandelkern, R. E. Olmstead, J. Jou, E. Tiongson, V. Allen, D. Scheibal, E. D. London, J. R. Monterosso, S. T. Tiffany, A. Korb, J. J. Gan, and M. S. Cohen, "Neural substrates of resisting craving during cigarette cue exposure.," *Biol. Psychiatry*, vol. 62, no. 6, pp. 642–51, Sep. 2007.

[25] H. W. Chase, S. B. Eickhoff, A. R. Laird, and L. Hogarth, “The neural basis of drug stimulus processing and craving: an activation likelihood estimation meta-analysis.,” *Biol. Psychiatry*, vol. 70, no. 8, pp. 785–93, Oct. 2011.

[26] X. Xu, J. Wang, A. Aron, W. Lei, J. L. Westmaas, and X. Weng, “Intense passionate love attenuates cigarette cue-reactivity in nicotine-deprived smokers: an FMRI study.,” *PLoS One*, vol. 7, no. 7, Jan. 2012.

[27] J. H. Lee, Y. Lim, B. K. Wiederhold, and S. J. Graham, “A functional magnetic resonance imaging (FMRI) study of cue-induced smoking craving in virtual environments.,” *Appl. Psychophysiol. Biofeedback*, vol. 30, no. 3, pp. 195–204, Sep. 2005.

[28] R. C. deCharms, F. Maeda, G. H. Glover, D. Ludlow, J. M. Pauly, D. Soneji, J. D. E. Gabrieli, and S. C. Mackey, “Control over brain activation and pain learned by using real-time functional MRI,” *PNAS*, vol. 102, no. 51, pp. 18626-18631, 2005.

CHAPTER 6

SUMMARY OF CONTRIBUTIONS AND FUTURE WORK

6.1 Novel contributions

- Implemented temporal brain state predictor for

  a) Classification using AVAST images –

  AVAST outperforms ASL while preserving all of its desirable properties that are preferred over BOLD for neural processes with sustained activation periods.

  b) Pain vs rest in TMD patients vs healthy controls –

  Significant differences were observed in the pain vs rest classifier in patients and controls for pain applied to thumb as well as face.

  c) Graded fMRI activation –

  Support vector regression models were trained to accurately predict fMRI activation on a continuous scale in motor activation and craving modulation.

- Subject categorization into healthy controls and autistic patients –

  Resting state functional connectivity maps were used to categorize subjects into controls and patients with high accuracy; significant weight clusters were found in the default mode network area.

- Implemented multi-subject machine learning classifier for craving population

  A population-wide shared component and subject-specific individual components of craving related brain activation were identified in the significant weight clusters.

- Set up real-time fMRI neurofeedback

  MATLAB was used to tailor the craving paradigm for each subject and then display images as well as classifier output of the machine learning algorithm sent over a real-time UDP communication channel. This setup was used to enable motor activation neurofeedback and craving modulation.

6.2 Summary and future work

In this thesis, we have explored various methods and applications of machine learning techniques in different functional MRI experiments. These algorithms have been successfully applied to perform various tasks such as brain-state classification, categorization of subjects and also shown to characterize and predict graded activation using regression methods. Although we demonstrate compelling results in the dissertation, these findings are by no means exhaustive. There are several topics that remain to be explored. Following is a summary of specific contributions and possible future work.

In chapter 3, we have proposed the use of an arterial cerebral blood volume (aCBV) weighted acquisition technique called AVAST in a support vector machine (SVM) based temporal brain-state classification experiment. This technique exhibits results that are comparable to BOLD and superior to perfusion-weighted ASL while retaining desirable characteristics of ASL imaging. AVAST involves calibration scans to tailor the best timing parameters for each subject. Further exploration could include acquiring images using optimal as well as sub-optimal timing parameters for each subject and then comparing the classification accuracy of the brain-state classifier to further emphasize the significance of individually tailored AVAST acquisitions. Furthermore, in this study, we have investigated the efficacy of this technique on robust visual-motor activation but this can be easily extended to brain-state

classification of other more complex processes including craving states in nicotine dependent subjects since craving processes exhibit a long sustained period of activation.

In the experiment involving brain-state classification of pain vs rest, the experimental design involves controlled application of two different levels of pain. Thus, this problem can be formulated for regression or multi-class SVM approach. The weight maps generated in these cases could throw light on important neural substrates of pain regulation in healthy subjects and subjects suffering with temporomandibular disorders (TMD).

Support vector regression (SVR) was used to predict graded fMRI activation on a continuous scale in motor activation. This is further extended to use with craving activity with subjects' self-reported craving measures as input labels and then predicting their craving measure on a continuous scale during testing phase. This can be extended to many other neural processes, especially, those that involve graded activations.

In chapter 4, we used resting state scans and computed correlation maps using a seed in the posterior cingulate cortex (PCC). The resting state functional connectivity (RSFC) maps obtained were used to distinguish healthy subjects from subjects with Asperger's disorder. In this study, we use SVM classification for the categorization of subjects. Instead of using a binary label based on disease state, it is possible to use a continuous label and perform regression on these maps. This is especially possible with autistic subjects since the diagnosis of autism is based on a wide spectrum and each diagnosed subject is given a score that defines the severity of their autistic traits. This score can be used as a continuous label input to the support vector regression (SVR) algorithm. Thus SVR output would characterize where on the autism spectrum the subject lies. This measure could be used as an additional guideline by clinicians and not the only tool to decide the severity of autism.

In chapter 5, we introduced a novel multi-subject machine learning approach to build a more informed model using multiple subjects' data at once. This technique uses multi-task learning methods with functional MRI data where each individual subject's brain activation is treated as a separate task. The model formed using this method, can be subdivided into a population-wide component and an individual-specific component. A detailed analysis of the optimal value of coupling parameter chosen for each subject could help characterize how similar they are to the rest of the nicotine dependent subjects. This could then be correlated with other phenotypic details such as motivation to smoke, age, gender, etc. to construct better models of smoking behavior and build insight.

We have described a real-time neurofeedback setup that we developed to enable self-regulation of nicotine craving. Future work could include studying the possible reinforcing effects of providing placebo feedback. This would involve presenting the subject with an ideal feedback regardless of their actual brain activation. The knowledge of being able to control their craving-related brain activation could be empowering and assist them in being able to quit. The real-time neurofeedback setup can also be used to study several different neural processes, especially those processes that can be individually regulated or modulated.