Automated Detection and Quantification of Pain Using Electroencephalography

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Vishal Vijayakumar

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

Professor Bin He, Advisor

May 2018

Acknowledgements

First and foremost, I would like to thank my advisor Dr. Bin He for his continued guidance, support and encouragement for my work. His timely advice and direction have enabled me to push past many plateaus during my years as his PhD student. I would also like to thank my committee members Dr. Keshab Parhi, Dr. Mehmet Akcakaya, and Dr. Tay Netoff for their support as members of my PhD committee. In addition, I would like to thank the Graduate School, the Department of Electrical and Computer Engineering, the Department of Biomedical Engineering at the University of Minnesota, as well as the National Institutes of Health and the National Science Foundation for their financial support during my years of research.

I would like to thank my lab mates Michelle Case and Sina Shirinpour for helping me design and execute the experiments to collect all of the invaluable data for my research, as well as their endless patience and advice during my sudden (and sometimes redundant) brainstorming sessions. I would also like to thank my lab mates Abbas, Brad, Chris, Daniel, Jeet, Ting, Haiteng, Yicun, Shuai, Amir, Bryan, Jake, Kai and Steve for their friendship, illuminating discussions about research and the road ahead for all of us. I want to thank Krithika Rai for enabling and encouraging me to pursue my PhD degree in the face of overwhelming odds, for all the walks in

and around the U of M that always ended too soon, conversations over nothing and everything, breakfasts at Al's, and especially for being there for me when it meant the most.

Most of all, I would like to thank my parents for their unwavering love, supporting me through their words, actions, and prayers, and who have sacrificed more for me than I will ever know.

Abstract

Effective pain assessment and management strategies are needed to better manage pain. In addition to self-report, an objective pain assessment system to detect, quantify, and track the intensity of pain reduces the uncertainty of treatment outcome and provides a reliable benchmark for longitudinal evaluation of pain therapies. As the brain is the primary generator of the pain response, neuroimaging methods hold the most promise for unravelling the mystery of pain. To analyze the temporal interactions of brain regions known to be active during pain, electroencephalography (EEG) is slowly gaining traction as a reliable tool for characterizing the feeling of pain. This work presents the development of robust and accurate machine learning algorithms on neuroimaging data using EEG to detect and quantify tonic thermal pain, and look into the temporal pain response to interpret the neuroscientific significance of the findings.

First, a random forest model was trained to predict pain scores using time-frequency wavelet representations of independent components obtained from EEG data. The frequency spectrum was used to determine the frequency band most informative for pain quantification. EEG source imaging was utilized to discriminate between the spatial correlates of a pain signature shared across subjects and a local pain signature unique to subjects. The mean classification accuracy for predicting

pain on an independent test subject for a range of 1-10 is 89.45%, highest among existing state of the art quantification algorithms for EEG. The gamma band is the most important to both inter-subject and intra-subject classification accuracy. A pain signature common across subjects localized to the dorsal mid-cingulate cortex while one unique to a subject predominantly localized to the anterior cingulate cortex, leading to a discussion of the observed differences between the two types of responses. The robustness and generalizability of the classifier is demonstrated, and there are two co-occurring neurological pain responses for the same stimulus which are relevant for quantification. These results demonstrate the potential of this tool to be used clinically to help improve chronic pain treatment, and establish spatial biomarkers for future pain-related studies using EEG.

Next, a temporally pain-specific biomarker using EEG was developed. Developing pain-specific biomarkers is challenging due to the spatial proximity of pain processing regions to other regions of the brain required for attention and other cognitive tasks, necessitating an investigation into their temporal behavior. The functional behavior of brief periods of global quasi-stable topographies of scalp electrical potentials in EEG, called microstates, was studied across 27 subjects to evaluate their specificity to pain compared to rest and two non-rest conditions evoking similar responses, such as non-painful heat and an N-back attention task. The distributions of microstate duration, occurrence, and coverage were tested for

significant differences between rest, pain, warmth, and task. In addition, a multifractal analysis on the microstate time series was performed to characterize temporal microstate interactions across several time scales and differences in these interactions across all conditions. It was found that two microstates related to emotional salience and attentional functional networks showed reduced duration and increased occurrence during tasks compared to rest. First and second order multifractal parameters extracted from the temporal microstate sequence showed that microstate interactions during pain were significantly more stable across time scales compared to non-painful conditions, but significantly more chaotic compared to resting state. For the first time, it has been shown that EEG is capable of detecting a pain-specific spatiotemporal response in the brain. This will help build toward an objective method to detect and quantify pain by being able to identify unique temporal patterns related strictly to pain.

Finally, a fast pain detection algorithm using deep learning techniques was constructed. A generalization of the temporal microstate sequence was used, where non-orthogonal temporal dependencies between instantaneous EEG spatial topography maps were characterized by the distribution of a dictionary of microstate fits per map. The deep learning network was developed as a tree classifier, with a binary gated recurrent neural network forming each branch. Each branch was trained to differentiate between pain and a non-painful condition in order to increase the

specificity of the final algorithm to pain. The algorithm was trained and validated on data collected from 22 subjects and tested on 5 subjects. The resulting algorithm improved on the state of the art by 14%, scoring 90.67% in terms of specificity to various levels of pain, compared to both attentions to non-painful heat stimuli and particularly demanding memory-based N-back tasks. Stacking this deep learning pain detection algorithm on top of the pain quantification algorithm developed earlier to produce a two-stage unified pain detection algorithm showed a 10% improvement in terms of F-score over the state of the art in pain quantification algorithms. This is an encouraging step forward in developing a clinically feasible tool that can detect, record, quantify and longitudinally compare the intensities of pain in patients to better aid the development of effective therapies to manage pain.

Contents

List of Figures

Chapter 1

Introduction

1.1 Pain

1.1.1 Overview

Pain is a conscious experience, a result of the brain processing noxious input influenced by memories, emotional, pathological, genetic and cognitive factors. The relationship between the intensity of the noxious stimulus and the level of pain is non-linear, and the behavioral response is often modulated by the situation experienced by the subject. Hence, pain is defined by the International Association for the Study of Pain (IASP) as, "an unpleasant sensory or emotional experience associated with actual or potential tissue damage or described in terms of such damage". Figure 1.1 illustrates the many factors that have been known to modulate the feeling of pain [1], resulting in many theories about the flow of control in the central and peripheral nervous systems influencing the degree of pain felt in humans and animals [2], [3].

Chronic pain is a clinical condition where persistent pain is experienced for longer than three months, leading to a significant loss in quality of life for the patient,
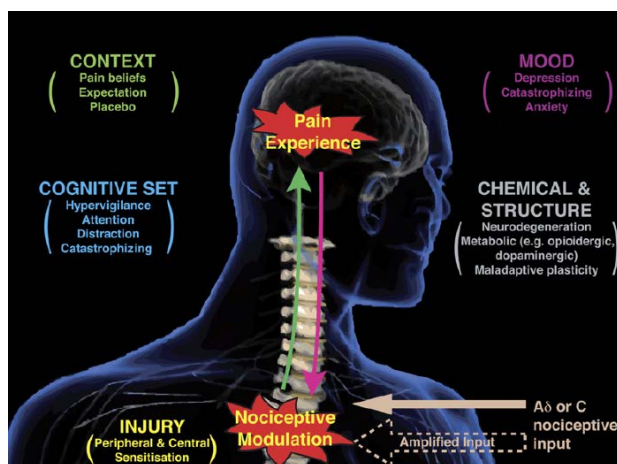
Figure 1.1: Schematic illustrating the main factors that influence nociceptive inputs to affect pain perception

hindering their ability to work and function in society. The most common medication prescribed for chronic pain are opioid analgesics to pursue more aggressive pain management, which the clinician prescribes in response to the self-reported pain score given by the patient. Unfortunately, the addictive nature of this drug has resulted in increasing trends in overdose-related deaths in the United States [4], [5]. To gauge a patient's pain state, clinicians rely on a set of answers to predetermined survey questions, the most common being the visual-to-analog scale (VAS). This scale usually requires patients to submit self-reported pain scores on a scale of 0 (no pain) to 10 (worst pain imaginable). After a prescribed dosage of medications on the first visit, follow-up visits are required to evaluate the success of medication, which again relies on a subjective pain score.

However, patients have difficulties in consistently providing a reliable estimate of their pain, as it is easily influenced by other factors affecting their mental state. This problem is compounded by the fact that patients have no reliable baseline to differentiate between pain scores between successive visits due to the fast forgetting factor of the feeling of pain [6]. The problem is two-fold: overdosing the patient can lead to an addiction to opioids and prescribing lower-than optimal dosages does little to alleviate pain. Hence, the lack of a reliable visit-to-visit baseline pain score creates uncertainty in the success rate of a treatment, leading to massive expenditures in the healthcare industry [7]. Developing an objective pain detection and quantification algorithm that can (1) provide consistent, reliable estimates of pain, and (2) that can record and compare the progression of pain in response to therapy over time can provide a possible solution to many of the problems enumerated above, forming the motivation for this thesis. As the brain is the primary generator of the pain response, this work attempts to achieve the goal of detecting and quantifying pain by analyzing relevant neurophysiological patterns in neuroimaging data collected from human subjects.

1.1.2 Neurophysiological Correlates of Pain

The nociceptive information is transmitted from the site of stress through the fast Aδ and the slower C nerve fibers, ascending to the thalamus through the spinal tract into the medulla and the brainstem to form short-range connections to the insula

and putamen. From this site, many research studies have demonstrated nociceptive information processed in the primary and secondary somatosensory cortices, the anterior cingulate cortex, the posterior parietal cortex, the premotor and the primary motor cortices [8]. From these studies, it is clear that several regions in the brain are more or less active in pain processing, depending on the interplay between factors modulating the pain sensation (mood, cognition etc.). Hence, it is agreed that there is no one primary region in the brain responsible for pain. Rather, these regions form a loosely connected spatiotemporal system, collectively called the *pain neuromatrix* [9]. However, the specificity of these regions to pain has been highly contested [10], [11]. Hence, recent neuroimaging research has shifted away from independent analysis of the activity of these spatially distinct regions towards the interaction between these regions in the temporal domain [12], in an effort to understand the integration of the complex experiences that form the feeling of pain.

1.2 Electrophysiological Imaging for Pain

Noninvasive imaging using electroencephalography (EEG) has become a popular method of electrophysiological imaging in recent years. Since 1929, when the German psychiatrist Hans Berger [13] described recording electrophysiological signals from the surface of the scalps of bald men, the EEG has provided an effective means of looking at brain activity. Some of the main applications of EEG include
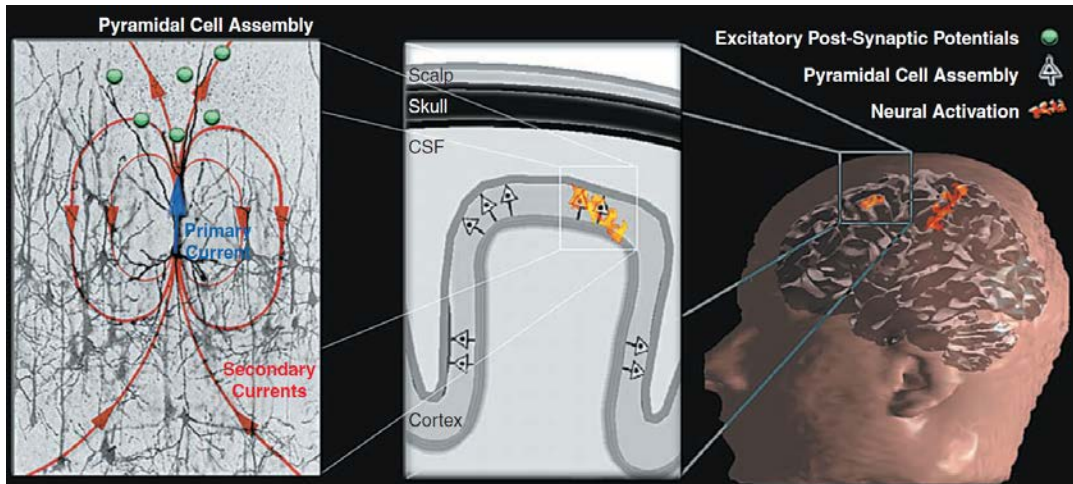
Figure 1.2: Neurophysiological mechanism of acquiring EEG signals

differential diagnosis and detection of epilepsy [14], sleep analysis [15], and Brain-Computer Interfaces [16] among many others. For imaging pain, researchers have been able to identify correlates of evoked pain in the continuous EEG [8], [17]–[19]. Due to the high temporal resolution of EEG, the most common method of characterization of EEG data for pain research in the neuroscientific field has been in the frequency domain. The standard frequency bands used are the delta band (2-4 Hz), the theta band (4-8 Hz), the alpha band (8-12 Hz), the beta band (12-30 Hz) and the gamma band (30-80 Hz).

1.2.1 Neurophysiological Basis for EEG signals

Figure 1.2 [20] shows the mechanism of generation of EEG signals. The gray matter of the neocortex in the human brain consists of six layers, with layer 1 being

closest to the cortical surface and layer 6 adjacent to the white matter, consisting only of axons that connect to different parts of the brain. In addition to the interneurons that form short-range connections, the neocortical grey matter also contains pyramidal neurons that send long axons to other parts of the CNS, where they form excitatory connections. EEG signals are hypothesized to be generated primarily by the pyramidal neurons in layer 5, close to the white matter, and extend to the cortical surface through long dendritic trees. These dendrites form electrical dipoles either through input from the superficial layers, or through layer 4. When activated simultaneously, these pyramidal neurons form dipoles adding up to a large signal that can be detected in a non-invasive fashion.

However, this does not detect intracellular currents directly. It measures the potential difference generated between the inner and outer cell membranes of the apical, non-depolarized parts of the dendrites, which leads to extracellular current flow. This current flowing through the volume conductor in loops, at various distances from the generating neuron, generates potential differences large enough to be measured by sets of electrodes placed on the scalp.

1.2.2 Advantages of EEG for Brain Imaging

EEG is a technique that is dependent on data from summed up regional neurological activity, which is a more "direct" method of measuring brain response

to stimuli, and can be effectively compared to invasive neuroimaging techniques [21]. As the temporal resolution of EEG is at the millisecond level, the dynamics of pain can be effectively characterized to uncover network activity. EEG equipment is portable, affordable and convenient to set up and use for the practitioner. The use of EEG caps allows the subject to be mobile due to the placement of a reference electrode on the cap, and the electrodes can be rapidly set in place consistently. Hence, these advantages demonstrate the potential of EEG to be used as a tool in the clinic for pain quantification and characterization, over conventional neuroimaging methods for pain research such as fMRI.

1.2.3 Limitations of EEG for Pain Research

As a non-invasive neuroimaging method, EEG does not measure signals at the cellular level, as its spatial resolution is of the order of 1 cm. Although many effective source imaging methods for EEG have been developed to address this issue [22], there is no unique solution, especially in problems where there are no *a priori* hypotheses about active regions of the brain. In addition, there is no true reference site on the body for measuring EEG signals. Electric potentials cannot be measured in absolute terms, but only as potential differences between two electrodes. Hence, placement of the reference electrode is critical for EEG recordings. Despite these limitations, temporal analysis of whole brain data has been shown to yield good

results for pain quantification [23]–[25], further adding to the potential of its effectiveness as a clinical diagnostic tool.

1.3 Pattern Recognition for Establishing Pain-Related Brain Biomarkers

The brain is the primary generator of the sensation of pain. To characterize it, a fundamental understanding of the brain mechanisms while processing pain is required before any meaningful treatment can be researched. To this end, researchers have been using neuroimaging techniques such as fMRI, PET, EEG and MEG for pain research over the last thirty years. Various sub-regions of the brain have been implicated and well-studied in pain processing, both for acute and chronic pain [8], [19], [26]. However, these studies focus on the brain mapping approach, which seeks to identify those brain regions responsible for certain psychological/somatosensory stimuli. The outcome of such studies is usually a statistical map of brain responses that highlight regions of interest. Such approaches subject information from each sensor (or voxel) to an independent analysis, attempting to characterize temporal activity over time to test against a null hypothesis of brain activity pertinent to a behavioral condition.

Recent advances in pain research have been in translating these basic findings into having useful brain biomarkers that help predict pain in individuals, determine
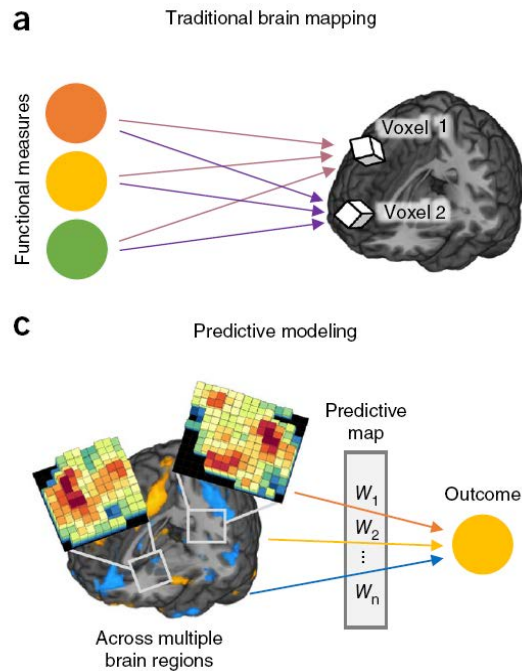
Figure 1.3: From brain mapping to brain mechanisms

the neurophysiological type of pain observed, and characterize its connections with other brain systems related to mediating fear, avoidance and depression. This approach inverts the traditional analysis, where instead of treating activity at the sensor (or voxel) as the outcome, it is treated as a characteristic used to predict/analyze a clinically relevant outcome such as pain intensity [27], duration [28], time of occurrence etc. Such neurophysiological data is often combined with other measurements such as heart rate, skin conductivity and eye-blinks in order to better predict the outcome. Hence, modern pain research has morphed into the search

for the brain biomarkers most strongly predictive of the desired diagnostic category or functional outcome, as shown in Figure 1.3 [29]. Efforts have also been made towards generalizing this approach to finding biomarkers consistent across subjects that are primarily responsible for regulating the desired outcome.

Hence, the brain biomarker approach to pain research used in this thesis is about establishing robust and descriptive brain activity in response to a particular stimulus category. However, such patterns of activity are not visible to the human eye. Automated pattern recognition algorithms, commonly falling under the umbrella of Multivariate Pattern Analysis (MVPA) techniques, are used to find and characterize these biomarkers. Out of all known algorithms, machine-learning techniques provide very powerful ways of identifying relevant brain regions and patterns of brain activity by training on an exhaustive set of labelled examples. The performance of these trained models can be evaluated by testing them on an independently collected test dataset.

1.4 Machine Learning on Neurological Data

Machine learning can be interpreted in the context of neuroscience as a tool that learns to decode patterns from datasets with hundreds to millions of variables. This approach tests whether individual samples can be classified into their respective class labels, after learning from a training set similar in distribution. Such models that

require a training dataset are called supervised models. In its simplest form, a supervised machine-learning model is a transformation from a space of data characteristics, commonly called features, to a class label associated with the datapoint. In the context of pain detection, the labels could be "pain" and "no pain". The weights applied on each feature during transformation is learned from the training data. The exponential growth of computing power and data availability in recent years has led to computer algorithms replicating almost human-level performances in object detection [30], tracking [31], and speech recognition [32] applications. Some of the most common machine learning algorithms used in pain research are the support vector machines (SVM) [18], [23], [25], Gaussian process classifiers [33], and LASSO-based regression techniques [27].

1.4.1 Supervised Learning: Training

In order to extract meaning from the training dataset, it has to be characterized in an effective manner. Such a transformation of the input data, called feature engineering, is a crucial part of training a supervised learning model, since the model is only as good as the representation of the data. Each sample can be represented as a n dimensional feature datapoint $\mathbf{x}_i$, where x_{ij} is a feature of the datapoint. The brain biomarker approach for future neuroscientific applications constrains these features to be interpretable. For example, in the case of EEG, this feature vector could be the activity in the frequency spectrum, where each spectral band has its own

neuroscientific significance. For each sample $\mathbf{x}_i$, there is a corresponding label y_i, which is usually discrete for classification. In the case of regression, the output would be a continuous, real-valued variable, which is the case for a model that attempts to predict the temperature on a given day. Hence, the goal of learning is to obtain the best possible model performance, subject to some additional constraints that are domain-specific. Mathematically, the estimate of a given model can be expressed as

$$\hat{y}_i = \mathbf{w}^\mathrm{T}\phi(\mathbf{x}_i) + b_i \tag{1.1}$$

In equation (1.1), the trainable weights $\mathbf{w}$ form the association between the datapoint and the class label. The kernel ϕ is a mapping of the input to a higher dimensional space where the features can be linearly weighted. The training phase can be expressed as an optimization problem that attempts to minimize the difference between the estimate of the output $\hat{Y}$ and the ground truth output Y, expressed as the loss function $L(\hat{Y}, Y, \Theta)$, and Θ is a set of hyperparameters for the supervised model that are tuned during the validation phase.

1.4.2 Supervised Learning: Validation

The validation phase is used to estimate the set of hyperparameters that best minimize the loss function. The validation set is usually similar in distribution to the training set and collected independently of the training set. The validation phase is used to prevent overfitting, where the hyperparameters can be tuned to fit the training

set very well, but the resulting model would be extremely sensitive to patterns of noise that are present outside this training set. The two quantities that are closely monitored during cross validation are called *bias* and *variance*. A model exhibiting high bias would have bad performance on the training set, typically a high loss and poor accuracy, which can be remedied by finding a better representation of the training data, or just taking more measurements for training. A model exhibiting high variance has overfitted on the training data, hence would perform poorly on the validation dataset. This is usually remedied by appropriately penalizing the loss function in a process called regularization. The best model would show a good trade-off between bias and variance.

1.5 Aim and Outline of this Thesis

The aim of this thesis was to investigate the spatiotemporal biomarkers implicated in pain processing over the last fifty years of pain research and develop machine-learning algorithms to learn how to effectively detect and quantify pain. Due to the rate of innovation in methods of EEG data acquisition such as wireless and dry electrodes [34], [35], developing a fast and effective pain detection and quantification tool using EEG would be advantageous and convenient for clinical use. This would help the clinician to make a more informed diagnosis and prescribe accurate dosages of opioids to manage pain better. In addition, establishing and

investigating patterns of activity in the brain possibly unique to pain would also accelerate the development of electrophysiological therapies for chronic pain management.

Due to the non-specific nature of spatially distinct brain regions to pain processing, an analysis is carried out in the temporal domain using EEG to identify and characterize the temporal pain response in healthy subjects, with investigations into the robustness and specificity of this pain signature across the subject population. New studies have been designed and executed to collect data corresponding to a consistent *relative* pain score collected across the subject pool, where the subjects rate their pain relative to their unique lower and upper bounds on the pain they can tolerate. The bulk of this thesis is divided into two parts: (a) quantifying pain given that there is pain in the first place, and (b) investigating the specificity of the temporal pain signature when considered with other confounding factors.

Chapter 2 presents a new study design and machine learning algorithm pipeline for quantifying tonic thermal pain. The study focuses on consistency in the pain rating rather than stimulus intensity to predict accurately the pain score of an unseen subject based on their continuous EEG data.

Chapter 3 analyzes limitations in current pain research. It details the steps to develop an effective spatiotemporal representation using EEG that has a distribution unique and specific to pain. This representation is derived from the temporal interactions between quasi-stable whole-brain scalp potential topographies called microstates. This chapter statistically investigates the pattern of activity of these microstates over several time scales in order to differentiate between pain and non-painful task-related conditions.

Chapter 4 presents deep learning architectures for fast and effective pain detection using recurrent neural networks. These architectures learn the pattern of temporal dependencies detailed in chapter 3 in order to detect the onset of pain while the subject is performing a non-painful attention-related task.

Finally, Chapter 5 presents the conclusions and the future steps that can be taken to extend this work.

Chapter 2

Quantifying Tonic Thermal Pain

2.1 Introduction

As pain is considered a subjective experience, the most prevalent means of measuring it has been through self-report. While there are established and effective treatments for acute pain [1], the high degree of variability exhibited by pain scores over a period of time [36] in chronic pain patients results in the prescription of inadequate pain management strategies, necessitating the need for objective methods for measuring pain.

Many recent studies have demonstrated the involvement of several spatially distinct cortical regions in pain processing. The most consistent regions are the anterior cingulate cortex (ACC), the primary (S1) and secondary (S2) somatosensory cortices, the prefrontal cortex (PFC), and the insular cortex [8], [19], [37]. Due to such spatially distinct regions showing concurrent activation during pain, traditional univariate approaches [38] are not suitable for capturing the spatio-temporal dynamics of the pain response. Hence, a branch of multi-variate approaches, collectively called Multivariate Pattern Analysis (MVPA) techniques [39], [40] has

been increasing in popularity over recent years [41]. Machine learning based approaches are a form of MVPA which focus on identifying patterns in data, and making predictions about future behavior of similar data [42]. Such approaches can potentially be used to solve problems associated with real-time brain decoding [43] and for selecting the best biomarkers for various acute and chronic medical conditions [44].

Electroencephalography (EEG) is a useful noninvasive tool for assessing pain responses due to its high temporal resolution, clinical convenience, and low cost of setup and maintenance [45]. Many studies have shown observable pain-related cortical activity using EEG [46]–[48]. Studies have also been carried out which effectively demonstrate the activity of certain frequency bands in the pain response, mainly the alpha band [49], [50], and more recently, the gamma band [51]–[53]. However, most of these studies have investigated the effects of phasic pain stimuli. To mimic responses triggered during chronic pain [54], there are a growing number of studies using tonic thermal stimuli [28], [55]–[58].

The aim of the present study was to build a robust and generalizable classification model for predicting various levels of tonic thermal pain from EEG data recorded in a group of subjects. This approach was used to train and test the performance of classifier models both across subjects and within a subject. Using the constructed model, the frequency bands that consistently contribute most to the

accuracy of the model were analyzed. Source imaging was used to localize the spatial correlates of the shared temporal pain signature that contributes most to the inter-subject classification accuracy, as well as the unique temporal pain signature that best informs the intra-subject classification accuracy.

2.2 Methods

2.2.1 Subjects

Twenty-five healthy subjects participated in this study with a median age of twenty-four years old, eleven of whom were female. The study was approved by the Institutional Review Board at the University of Minnesota and conducted in conformity with the Declaration of Helsinki. Subjects were screened so that none of the healthy volunteers had any history of mental illness. All subjects provided written informed consent before participating in the study.

2.2.2 Experimental Procedure

The experimental procedure consisted of four phases: pain threshold and tolerance determination, the tonic pain phase, and a control phase, performed on each subject in that order. A tonic thermal stimulus was delivered to subjects using a thermal stimulator (CHEPs, Medoc, Ramat Yishai, Israel). A custom-built analog

pain-rating device was used to record pain scores. Subjects could rate pain by indicating the level of pain intensity using the device and aided by a color bar with a gradient from green (indicating no pain and non-painful heat) to red (indicating high pain). A line was placed on the color bar to indicate pain threshold, and subjects were asked to rate any painful thermal stimuli above the bar. Before beginning the experiment, subjects were given time to practice with the device by rating thermal stimuli delivered to their left forearm. In this training session, subjects experienced different levels of the thermal stimulus ranging from non-painful heat to moderately high pain temperatures. This practice allowed subjects to accurately determine how to rate their pain on the provided color bar.

The phases of the experimental procedure are illustrated in Figure 2.1. To determine pain threshold, the heat delivered by the thermode was increased at a constant rate of 0.5˚C/second with a safety limit of 50˚C. Seven trials were conducted with an inter-trial interval of 10 seconds. During this phase, the subjects were asked to pay attention to the stimulus and use a response unit to indicate the instant the stimulus changed from heat to pain, after which the temperature was brought back to baseline at a rate of 5˚C/second. The rate of increase of temperature for tolerance testing was the same as that for the threshold test, with the same number of trials.

Experimental Phases
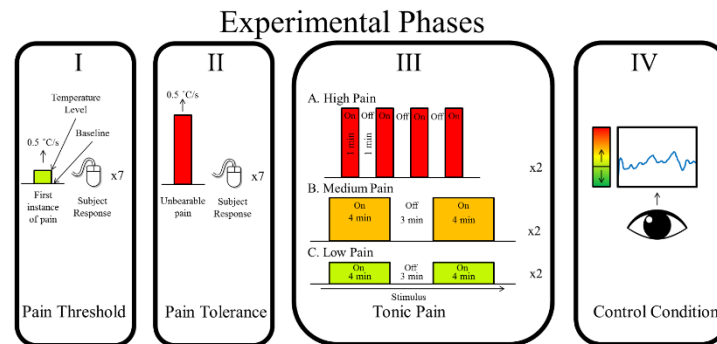
Figure 2.1: Stages of the experimental protocol

However, for the tolerance test, subjects were instructed to continuously rate their pain with the analog rating device using their left hand, and to use the response unit to indicate the temperature at which the painful stimulus became intolerable. The threshold and tolerance tests were used to determine the minimum and maximum temperatures to be administered to the subject during the tonic pain condition to ensure the desired range of pain scores was obtained for each subject and to make sure no one was administered an unbearable pain stimulus. The thermode was placed on the ventral left forearm for both the tests, while the right arm remained motionless.

The tonic pain phase was split into three conditions: low pain, medium pain, and high pain. The thermode was shifted to the right ventral forearm for tonic pain trials, which subjects were asked to keep motionless while they rated their pain with the device using their left hand. The low and medium pain conditions consisted of two trials of a constant temperature lasting four minutes each, with a resting period of

three minutes between each trial. Due to the intensity of the high pain condition, the trials were shortened to four trials of one minute each, with a resting period of one minute between trials. During rest, subjects were instructed to either stare at a fixed point on the screen, or observe a silent video clip of an aquarium to constrain their streams of consciousness. The subjects rated their pain from a predetermined temperature based on the threshold and tolerance responses using the analog device and the color bar. For the low pain condition, the subject was administered heat a degree higher than the median threshold response of the seven trials. During high pain, the administered temperature was half a degree lower than the median tolerance response of the seven trials. The temperature administered was constant to eliminate the possibility of saliency or awareness of temperature changes being reflected in the EEG recording. Subjects were not aware of the stimulus paradigm so that there was no inherent bias on how they would rate their pain. The only instruction given to the subjects during this phase was to try to accurately rate their pain intensity based on the stimulus received.

The visual control phase was used to account for the visual, attentional and motor components of all the trials to ensure that EEG components from these sources during the tonic pain phase do not adversely influence the machine learning algorithm [26], [28], [59]. During this phase, subjects were asked to use the pain rating system to rate the height of a continuously shifting point displayed on the

screen, while the temperature of the thermode was maintained at body temperature (37°C). The height of the point was determined by the temporally inverted pain rating of one of the previous trials chosen at random.

2.2.3   EEG Recording and Pre-processing

During all phases in the procedure, EEG data were recorded at 1 kHz using 64 channel Neuroscan system (Compumedics, Charlotte, NC). The impedance was kept below 5KΩ for all conditions. The pain rating was collected at a sampling rate of 10Hz.

During the analysis of each pain trial, the initial and final 10 seconds of EEG data were discounted to eliminate portions of the EEG signal with high variability and noise due to inconsistencies in initial and final sensations caused by the stimulus. These variations were more likely to reflect awareness of a drastic change in the stimulus rather than actual pain rating scores. The analysis was performed in MATLAB and Python using native toolboxes and those developed by EEGLAB [60].

The EEG signal was down sampled to 256 Hz and band-pass filtered between 1 Hz and 100 Hz to remove any muscle contraction artifacts and voltage drifts. Line noise at 60 Hz was filtered out using a notch filter. EEG was re-referenced to average reference. A modified version of the artifact subspace reconstruction (ASR) [61] was used to clean eye blink, muscle movement artifacts and reject bad time windows.

The rationale behind this algorithm is that the data collected using electrodes is highly correlated across channels, hence estimation of one noisy channel can be carried out by interpolating information from neighboring electrodes. Statistics for estimating noise thresholds are calculated from one minute of calibration data. A sliding window over the raw sensor is used to find subspaces using principal component analysis (PCA), with the highly variance components considered to be likely candidates for containing noise. Noisy segments are removed if activity is more than a few standard deviations away from the thresholds computed using calibration data. The signal is then reconstructed using the remaining PCs.

For calibration data $X_c^{M \times N}$, the PCs are estimated using the covariance matrix C. Upon computing the eigen-decomposition of $C = WLW^T$, the eigenvectors $W^{M \times Q}$ are then used to estimate the Q component activations $Y^{Q \times N} = W^T X_c$. For each component activation $y_q \in Y$, the robust mean m_q and standard deviation σ_q are computed and collected, to form a thresholding vector $z = m + c\sigma$, where c is a tunable parameter to set the upper limit on noise activity to be tolerated in the raw data. These thresholds are then projected to the component subspace of a given segment of the raw data given by $V = [v_1, v_2, \ldots, v_Q']$. The final projected thresholds are then $t(v_k) = \|zv_k\|_2^2$. If any segment exceeds the set threshold, it is removed in the component subspace. The remaining components are then recombined to obtain the cleaned sensor data.

Independent component analysis (ICA) [62] is then used to reject other noisy components, primarily eye movement, cardio-ballistic artifacts, and other bad electrode sources.

2.2.4 EEG Analysis

The EEG data corresponding to the pain and the control conditions were then concatenated along with the pain ratings, with the data acquired from control conditions assigned a pain score of one. Two minutes of resting state data were also appended to enable the model to discriminate between rest and pain states. The data were then concatenated across subjects, and subjected to full rank group ICA [63]. The rationale behind using this approach was that the pain sources found deep in the brain would have correlates found in a distributed subset of channels, rather than just one channel. As analysis of this multivariate dataset would be computationally intensive, ICA is used to project the data from multiple electrodes into a subspace where their activity can be represented by a single IC time course. The assumption made is that activity in brain sources are non-Gaussian in nature. ICA also facilities analysis the time course of each IC independently by definition. Coupled with the fact that muscle activity and other EEG artifacts are separated into their own components, this projection is useful for both improving the SNR and isolating the potential temporal pain signature into one or more components.

The ground truth obtained was labeled as follows: all the rest conditions, including the control conditions and non-painful heat, were assigned a pain score of one, while the continuous pain responses were discretized into '*n*' equal interval classes. The variable '*n*' is iterated from 2 to 10, 2 corresponding to a common 'pain' class, and 10 corresponding to 9 ordinal "pain" classes, and one rest class.

A continuous wavelet transform using the Gabor wavelet was used to extract the Time-Frequency Representations (TFR) [64], owing to its excellent tradeoff between temporal and frequency resolution among existing signal analysis methods (Sifuzzaman et al., 2009). Wavelet coefficients were computed for 60 scales, corresponding to a frequency range of 2-80 Hz. Choosing a smaller scale would result in more high frequency noise than actual signal, so a maximum of 80 Hz was chosen. This resulted in computing a feature vector of 60 scales per sample, while the number of samples skipped was made equal to half the width of the smallest wavelet scale used to analyze the EEG data. This was done to prevent redundancy in the feature space, while also analyzing all available data. All such time points computed were then used as data points to train and test the classifier model.

2.2.5 Pattern Analysis Using Random Forests

Random Forest (RF) models are an ensemble method of classification, where the fundamental classifier is a decision tree. A decision tree splits a dataset into

several leaf nodes in an attempt to "purify" the dataset at each node, where purification implies finding the best feature in the dataset to split a node to obtain leaf nodes which have a majority class [66]. In short, RF models are created by growing many unpruned decision trees trained on a random subset of the data and return a class for each data point that is the mode of all classes decided by all the trees. They are known for their high accuracies, efficiency on large datasets, and resistance to overtraining in the presence of sparsity as the number of trees increases by undersampling majority class labels [67]. This characteristic is relevant for the problem of pain quantification since the data corresponding to higher pain classes is relatively sparse. As the primary unit is a decision tree, the splitting rules for identifying the best feature used to split the training set forces the classifier to address all classes in an imbalanced dataset, making it a suitable choice for this study. The ensemble bootstrapping approach to quantification minimizes the trade-off between bias and variance. They also do not make any assumption about the distribution of the data or interactions between the features.

2.2.6 Performance Metrics for Quantification

Due to the non-uniform distribution of the discrete pain scores collected for this study, metrics other than classification accuracy were used to judge classifier performance during training and testing. The confusion matrices obtained during the

training and testing phases were used to calculate statistics such as Balanced

Classification Accuracy (BCA), F-measure (F) and the Matthew's Correlation

Coefficient (MCC). For a given class i in a multi-class setting consisting of N

possible classes, each of the metrics are defined below for a confusion matrix $C_{N \times N}$

as:

$$BCA = \frac{1}{N} \sum_{i=1}^{N} \frac{C_{ii}}{\sum_j C_{ji}} \qquad (2.1)$$

$$Precision, P_i = \frac{C_{ii}}{\sum_j C_{ji}} \qquad (2.2)$$

$$Recall, R_i = \frac{C_{ii}}{\sum_j C_{ij}} \qquad (2.3)$$

$$F_i = \frac{2 \cdot P_i \cdot R_i}{P_i + R_i} \qquad (2.4)$$

$$MCC = \frac{\sum_{k,l,m} C_{kk} C_{ml} - C_{lk} C_{km}}{\sqrt{\sum_k \left[\left(\sum_l C_{lk} \right) \left(\sum_{f \neq k,g} C_{gf} \right) \right]} \sqrt{\sum_k \left[\left(\sum_l C_{kl} \right) \left(\sum_{f \neq k,g} C_{fg} \right) \right]}} \qquad (2.5)$$

These measures were computed from confusion matrices calculated for each test

subject to ascertain the robustness of the classifier. The trend of the BCA, which is

the mean of the precision across all classes, is analyzed across all subjects for various

resolutions of the continuous pain score. The F-score is the harmonic mean of the

precision (positive predictive value) and the recall (sensitivity). The MCC in

equation (2.5) is a robust measure of the quality of the classifier when it is trained/tested on imbalanced classes. It can be interpreted as a measure of correlation between observed and predicted classes, and has a range of [-1, 1] where a score of 0.4 (70%) or higher indicates good agreement between the observed and predicted class labels.

2.2.7 Training and Testing Paradigms for Quantification

The efficacy of the quantification approach was tested both across subjects and within subjects. To test performance across subjects, Leave-One-Out Classification (LOOC) was used to train the classifier across 24 subjects and test on one subject. An illustration of the training and cross-validation pipeline for LOOC is shown in Figure 2.2 (I). The wavelet power spectrum of each time instant of a group IC, along with the corresponding discrete pain scores was used as data for a RF model.

The training data for the LOOC case is over a million data points, while the test set is about 100,000 data points. The goal was to find the model and the corresponding group IC that contained the best representation of the pain response. Each model was trained on two-thirds of the training data using various resolutions of the continuous pain score, and cross validation was performed on the remaining one-third of the training data where the accuracy of each model was returned. The model returning the best cross-validation performance metrics, along with the spatial
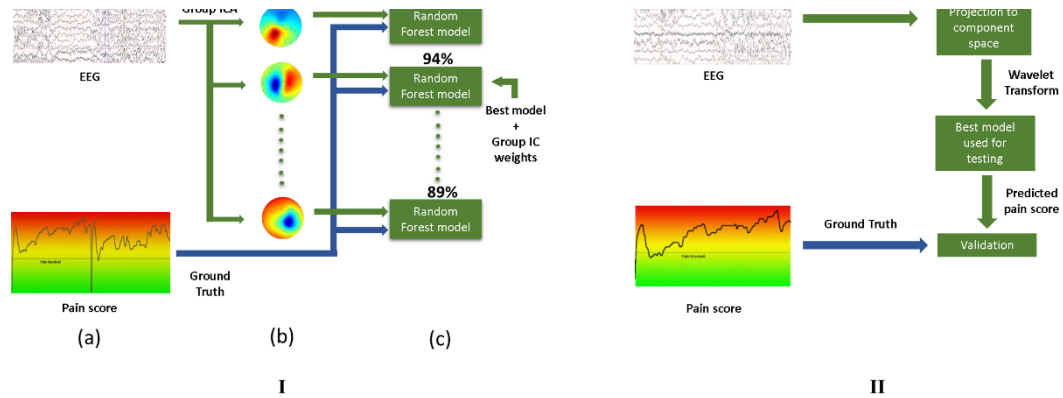
Figure 2.2: : Algorithm workflow. (I) Training paradigm. (a) Training EEG data, along with the ground truth is collected from all but one subjects. (b) Group ICA is used on concatenated pre-processed EEG data. (c) Random Forest models are trained on two-thirds of the training data for each component, and the best performing model along with the corresponding group IC weights are picked for testing. (II) Testing paradigm. The best model selected in I(c) is used to test performance of the prediction method on the unseen test subject

IC weights corresponding to the most discriminative group IC were used in the testing phase shown in Figure 2.2 (II). In this phase, the EEG data collected from the remaining test subject were projected into a new subspace using the weights of the most informative group IC obtained during cross-validation. The corresponding RF model was then used to estimate the discrete pain score for each time instant of the wavelet power spectrum calculated from the projected EEG data. Statistics of model performance were then computed using the ground truth pain score for that test subject. For evaluating within-subject performance metrics, the pipeline in Figure 2.2 (I) was used to train each subject, where the two-thirds of the wavelet scalogram

of an IC was used to train a specific model. The model with the best training performance was then tested with the remaining one-third of the corresponding IC time course.

To investigate within-subject classifier performance, two-thirds of the subject's EEG data were used to compute the ICs and individual RF models were trained on the corresponding wavelet time courses. The statistics of classifier performance (BCA, F-score, and MCC) were used to select the most informative IC using ten-fold cross validation. The corresponding spatial map was then used to project the remaining one-third of the EEG data to the component subspace, and the wavelet time course of this test data was used to test classifier performance. This process was repeated thrice to utilize all the EEG data for that subject.

In addition to predicting the pain score, analysis was also performed on the frequency spectrum on the most accurate test ICs for each subject to determine the frequency band most relevant to classification. Consider the feature set $F = \{f_1, f_2, \ldots, f_M\}$ where there are M features per data point. Let the decision set across all trees in the RF classifier be denoted as $D = \{d_1, d_2, \ldots, d_N\}$, where d_i is the pain score predicted by the classifier, and N is the number of decision trees in the forest. To determine the importance of a feature j within the testing set, the values of this feature were then randomly permuted within its range across all data points in the test set. The modified test set is then projected down the unchanged decision tree.

Let the new decisions be denoted as $D^j = \{d_1^j, d_2^j, ..., d_N^j\}$. The number of votes cast for the correct class is subtracted from that of the unperturbed decision tree. The magnitude of this deviation is the raw importance score, denoted by $R^j = \{r_1^j, r_2^j, ..., r_N^j\}$, where:

$$r_i^j = d_i^j \oplus d_i, j = 1 \dots M, i = 1 \dots N \qquad (2.6)$$

The final score is then computed by dividing the mean of the set R^j by its standard deviation. The feature importance score C for a feature j is thus given as

$$C^j = \frac{E(R^j)}{\sigma(R^j)} \qquad (2.7)$$

Hence, the final set $C = \{C^1, C^2, ..., C^M\}$ provides the feature importance scores for every feature independent of the other in the test dataset. The features with the highest scores are found to contribute most to the classification accuracy.

2.2.8 Source Imaging of the Temporal Pain Response

When trained across subjects, a basic, shared spatiotemporal pain signature that is common across that subject pool is more likely to be learned by the model. Hence, using this inter-subject training approach, the most informative group IC would tend to best exhibit the generalized pain signature for that subject pool. However, it can be hypothesized that any model trained and tested on a single subject is more likely

to pick an IC, which encodes his/her unique pain signature. Such a unique pain response could exhibit different spatial characteristics compared to the shared signature obtained using the previous approach. Source localization was used as a tool to investigate such spatial differences between the shared and unique pain responses for each subject using the most discriminative spatial IC maps obtained previously. As a spatial IC map assigns weights to each electrode proportional to its activity in the scalp EEG, source analysis on these maps could reveal the underlying sources of activity involved with pain. For this purpose, Principal Component Analysis (PCA) was performed on the spatial IC maps collected across subjects within each group (shared or unique). This accounts for the variability between the spatial maps across subjects, which would make source analysis more robust. Using PCA on the scalp maps would find the set of components that best represents the distribution of scalp electrode activity within the subject pool. Physiologically, the first component would encode the best spatial representation of the most consistent scalp electrode activity across subjects in a group. Subsequent decorrelated components would best represent sets of electrodes that are active less frequently across subjects. Source localization on the components that best explain the spatial distribution of scalp EEG would hence reveal the most consistent sources across the subject pool.

The analysis was performed using the sLORETA algorithm [68] to obtain the current density sources, using the Colin27 realistic standard head model for all subjects. A three-layer boundary element method (BEM) [69], [70] was used to separate the head volume conductor into brain, skull and scalp with the conductivity ratio of 0.33, 0.0165, and 0.33 S/m, respectively [71], [72]. Source locations were restricted to gray matter.

2.3 Results

2.3.1 Results of Classification Analysis

The statistics for the LOOC paradigm performed across subjects are presented in Figure 2.3. The horizontal axis shows the number of intervals that the continuous pain score is discretized into, while the vertical axis shows the classification accuracy. The BCA for 2-way classification is 95.33±0.6% and is 89.45±1.05% for 10-way classification, where chance BCA is 10%. While the slope indicates expected deterioration in the performance of the classifier as the ground truth intervals are made finer, the low rate of decrease demonstrates the scalability of the classifier to finer resolutions of continuous pain scores.

In addition to BCA calculated for different resolutions of the pain score across subjects, the confusion matrix computed for worst case (10-way) classification for

one test subject is shown in Figure 2.4. Most misclassifications occurred within the neighborhood of the true class label, and on average, 73% of the total misclassifications occurred within the lower half of the pain scale and 62% of all errors occurred within a 2-neighborhood distance from the true class. Hence, it is seen that misclassifications are more prevalent in the lower half of the pain scale than the upper half, and most of the wrongly predicted scores are about the true classes. From a clinical perspective, this is less of a risk compared to a larger misclassification rate on higher pain scores, and across larger pain intervals.

For a test subject, the BCA of each component is shown in Figure 2.5. The components are sorted in descending order of BCA performance from left to right for 10-way classification, where the dotted red line corresponds to chance BCA. The spatial topographies of the three best and worst performing components are also depicted. As described in [73], muscle artifacts show the strongest activity in pre-frontal and temporal electrodes, with spikes in higher frequency bands. It is seen from Figure 2.5 that ICA succeeds in isolating such type of activity. These ICs also perform consistently poorly across subjects for various resolutions of the pain score, showing that classification accuracy is not informed by any muscle activity. Alternatively, these ICs can be removed manually prior to classification to speed up computation time.

The performance metrics computed for the confusion matrices in the 10-way classification for both the cross-subjects and within-subject paradigms are shown in Figure 2.6. All performance metrics are higher when the classifier is trained and tested within subjects, as expected. The F-score for each subject is averaged across all classes. The F-score for both the inter-subject and the intra-subject paradigms (91.3±3.4% and 94.2±2.6%, respectively) is slightly higher than the BCA (89.45±1.07% and 93.26±1.49%, respectively). The MCC is 0.466 (73.3%) for the inter-subject paradigm, while it is 0.609 (80.46%) for the intra-subject paradigm, indicating strong agreement between the ground truth and predicted pain scores in the presence of class imbalance.

2.3.2   Frequency Band Contributions Towards Pain Prediction

In addition to predicting the pain score, analysis was also performed on the frequency spectrum on the most accurate test ICs for each subject to determine the frequency band most relevant to classification. For the best feature split at each node in each decision tree, the feature values were randomly permuted to assess the effect on the final class. The importance of the feature is directly proportional to the magnitude of deviation of the error produced by randomly permuting its values, as any association that existed between the feature and its corresponding class was

removed by random value permutation. The raw scores thus obtained were then divided by their standard deviation to obtain a normalized importance measure.

Estimation of frequency importance is performed on all subjects for ten-way classification. Figure 2.7 shows a boxplot depicting the normalized contribution scores for all subjects. It is seen that the gamma band contributes the most to classification accuracy, and that the alpha band has the highest variance across all subjects. The effect of each frequency band was also investigated in a paradigm where, during classification, the frequency information was ignored from one band at a time during training to observe its effect during testing. The results, illustrated in Figure 2.8 show that accuracy suffers the most when the gamma band is ignored, with a 10% decrease in performance compared to the original approach. However, it is interesting to note that accuracy suffers regardless of which frequency band is ignored, showing that all frequency bands seem to be important to pain score prediction.

2.3.3 Differences Between Global and Local Pain Signature

The percent of data explained by each component was used to assess how many principal components were examined for each group. A threshold of 90% was selected in both cases. Two components were assessed for the global group and three components were assessed for the local group.

For the inter-subject training paradigm, the most discriminative group IC was found for each training set, resulting in 25 group IC maps to be considered for PCA. Source imaging results on the first PC are shown in Figure 2.9. The most consistent source for the shared pain signature is found to be spread across the posterior Mid Cingulate Cortex (pMCC) and the dorsal Posterior Cingulate Cortex (dPCC), close to the cingulate sulcus. Subsequent non-noisy components show activation primarily in the MCC.

Similarly, for the intra-subject training paradigm, the most discriminative IC from each subject was collected for a total of 25 unique IC maps for PCA. Figure 2.10 shows source imaging results on the first PC for a similar analysis on unique ICs to determine sources most likely to encode the unique pain signature. The most consistent source of this pain signature is represented by activity in the ACC. The rest of the non-noisy components localize to the precuneus, the Posterior Cingulate Cortex (PCC) and the mid cingulate gyrus, respectively.

2.4 Discussion

2.4.1 Improvements over Previous Methods

A table comparing the current quantification method to existing state of the art methods is shown in Table 4.1. While such studies have had a fair amount of success,

most [18], [33], [74], [75] examined the efficacy of a two-class version of this problem (presence vs absence, or low pain vs high pain), while others [27], [76] pursued a regression-based approach. Those studies that pursued a multi-class approach [53], have broken it down to various two-class sub-problems with a vote-based approach to arrive at a pain score. In addition to the high computational time for training these models, such voting procedures would need testing data from both low-pain and high-pain trials to arrive at a single prediction score for a test subject, which may not be always clinically available. The current model substantially improves on the state of the art and is trained to accurately predict a maximum pain range of 1-10 based on any 2-3 second snippets of the time course of independent components derived from scalp-recorded EEG, which is clinically more convenient.

In addition, most of the past approaches using EEG have worked towards finding the most informative electrode i.e. the electrode where the prediction accuracy is the highest. While fairly accurate, such approaches may, in fact, be interpreted as univariate rather than multivariate [41] as a decision is made by processing data from one electrode at a time. In the end, this may not be the most informative due to the distributed nature of pain processing. By projecting EEG data into another subspace in which each component contains phase-locked information from multiple electrodes, group ICA isolates patterns recurring throughout the training data, which can generalize better for larger datasets as it favors selecting an IC with a more

consistent temporal pain signature. Even though Support Vector Machines (SVMs) are one of the most commonly used classifiers for pain prediction, issues with multi-class implementation and optimal parameter selection [77], [78] make RF decision trees a better alternative due to their scalability to large datasets and resistance to overfitting.

2.4.2   Spectral and Spatial Activity Most Important to Pain

The most informative IC found during testing for each subject was used to analyze inter-individual variability of the spatio-temporal pain signature in both the frequency and the topographical domains. It was found that 1) the gamma band (31-60 Hz) was the most important in discriminating between different pain scores, and the degree of importance was consistently high from two-way classification up to (and including) 10-way classification; 2) source imaging on PCA maps revealed that the most discriminative correlates of nociception were constrained primarily to the cingulate cortex, with spatial differences observed between the most informative global and local pain signatures.

The gamma band has been previously found to be important in the subjective perception of both acute and tonic pain. The sensory and attentional aspects of brain activity in response to acute pain were found to be most prevalent in the gamma band [28], [51], [52], [79], and the most obvious spatial signature was found near the

somatosensory cortex using scalp EEG data [51], [79], [80]. Some of those studies using both acute [52], [53] and tonic pain [28] also show increased gamma band activity in the medial pre-frontal cortex. Moreover, [79] showed that the predictive power of gamma band activity is independent of the rate of repetition of the stimulus. In this study, the high degree of importance of the gamma rhythm for pain quantification at both coarse and fine resolutions of the continuous pain score was observed across subjects for the first time in tonic thermal pain stimuli. Hence, it is likely that the gamma rhythm primarily encodes reactionary aspects of nociception for the entire duration of the stimulus. As previous studies investigating acute and tonic thermal pain have highlighted the gamma rhythm as well, this would point to the gamma rhythm being the most important biomarker for thermal pain.

Various sub-regions of the cingulate cortex have been frequently implicated in pain processing. The ACC has been associated with affective components of pain processing [81]–[83], with the pACC in particular showing elevated blood flow in response to cutaneous thermal stimulation [84]. Regions of the MCC have been correlated with emotion and avoidance in response to both acute and tonic cutaneous thermal stimulation [85]. The dPCC, in collusion with the pMCC seem to be involved in orienting the body towards noxious somatosensory stimuli and attention to the level of threat presented [86]. A summary of various regions activated in the cingulate cortex is presented in [83], where the most consistent spatial activations

during affective pain processing were observed in the pACC and the MCC without preference for sub-regions.

The current study corroborates the above findings and adds to them by showing that the best representation of a shared *spatiotemporal* pain signature can be observed in the dPCC and portions of the pMCC. This validates the hypothesis that such regions could be involved in early sensory pain processing that would tend to be more similar across populations. However, while the pACC and the aMCC seem to show more spatial consistency across subjects, the temporal pattern of the pain signatures in these regions is seen to be unique to the subject. This reinforces the notion that affective pain processing recruits similar areas of the brain across a population while the perception of pain is unique to each individual, explaining why such a temporal signature is not very informative in cross-subject quantification.

2.4.3 Salience

Salience, in the context of sensation and pain perception, is defined [87] to be the awareness of how much the stimulus intensity contrasts with the surrounding. It has been suggested in [10], [88] that the regions constituting the pain neuromatrix have been implicated in the processing of other auditory, visual, and tactile stimuli as well. Since non-nociceptive neurons have been observed to outnumber the nociceptive kind [89], it can be argued that any response predicted by this model in

fact, encodes sensory response, and may not provide information about pain processing.

In this study, the observation that regions in the cingulate cortex show activity in the presence of pain, illustrates their role in the processing of such stimuli irrespective of their exclusivity. This suggests that there is an evaluative component to pain that is distinct from that of other stimuli, even if such processes occur in tandem. Indeed, it has been shown that the dPCC is a central hub of the Default Mode Network (DMN), whose activity is related to internal processes and has reduced activity when tasks are performed [90], [91]. This inverse relationship of the DMN to attention might reflect attention to the task of rating pain, rather than pain processing, is contributing to pain prediction scores. To test if attention to the task was a factor in predicting pain scores, we compared the uncertainty of our classifier during the control task to rest, where no task was performed, shown in Figure 2.11. There was greater uncertainty during the control task, indicating attention does play a role; however, the reduced uncertainty during pain conditions shows the classifier is confident during the pain stimulus and attention to the task is not the only factor contributing to pain prediction. However, further studies need to be conducted as to the extent of the role of non-nociceptive stimuli in pain processing. Even though attempts are made to minimize effects of salience by keeping the stimulus intensity constant in this study, it is possible that over time the subject's attention drifted in

and out of pain recognition. It could be the dynamics of this process that is recognized by the model to strongly correlate with the pain score, given the importance of the dPCC. Although previous attempts have been made to model such dynamics of pain activity [6], [92], further investigations into the changes in the frequency importance scores over time can provide some more insight into modeling the temporal evolution of this signature.

2.4.4 Study Limitations

Pain is a sensation that is experienced and processed differently by everyone. While the initial shock of sensation is likely the same, the spatio-temporal correlates of affective and emotional processing are uncommon across individuals. While the ICs selected for the inter-subject case seem consistently accurate, there are significant spatial differences between those and the ICs chosen by the within-subject training paradigm. This shows that the uniqueness of the pain response topology diminishes as more subjects are considered for analysis, and this may account for a drop in the accuracy when more subjects are considered for training. With a high enough number of training subjects, the problem of pain prediction could potentially devolve into one of sensory discrimination localized to the caudal cingulate cortex.

Using mutual information in the RF model is an indirect measure of frequency importance, as it is not possible to identify a monotonic relationship between any of

the frequency bands with the pain score. To determine the extent of similarity between different frequency bands and the pain score at the electrode level, the best performing components were projected back to the scalp domain and the Welch power spectrum was computed for every second of the data to dynamically observe the changes in the frequency spectrum with time. Correlation analysis between any combination of frequency bands and the pain score did not yield results above significance, suggesting that the relationship between EEG data and the pain score is subtler than previously studied.

2.5 Conclusions

We have developed a novel noninvasive approach to the problem of pain quantification in the presence of tonic thermal pain and evaluated the excellent performance of our method in 25 human subjects with 89.45% accuracy in 10-way pain quantification. The temporal wavelet characteristics of independent components is used to characterize the EEG information corresponding to the pain score, and the most informative component is evaluated, along with observing its characteristics in the spatial and the frequency domains. The present approach outperforms existing state of the art methods for pain quantification by reliably predicting the pain score and demonstrating scalability of the approach for finer resolutions of a continuous pain scale. The present promising results suggest the generalizability of a biomarker

for pain across a population. The gamma band in EEG is found to be the most important predictor of pain for all resolutions of the pain score across inter-subject and intra-subject prediction paradigms. In addition, it is shown that the differences between the shared neural correlates of the pain score across populations and the unique signature at an individual level are observable, hence can be quantified in further explorations by extending these approaches; unraveling a potential pain response mechanism.
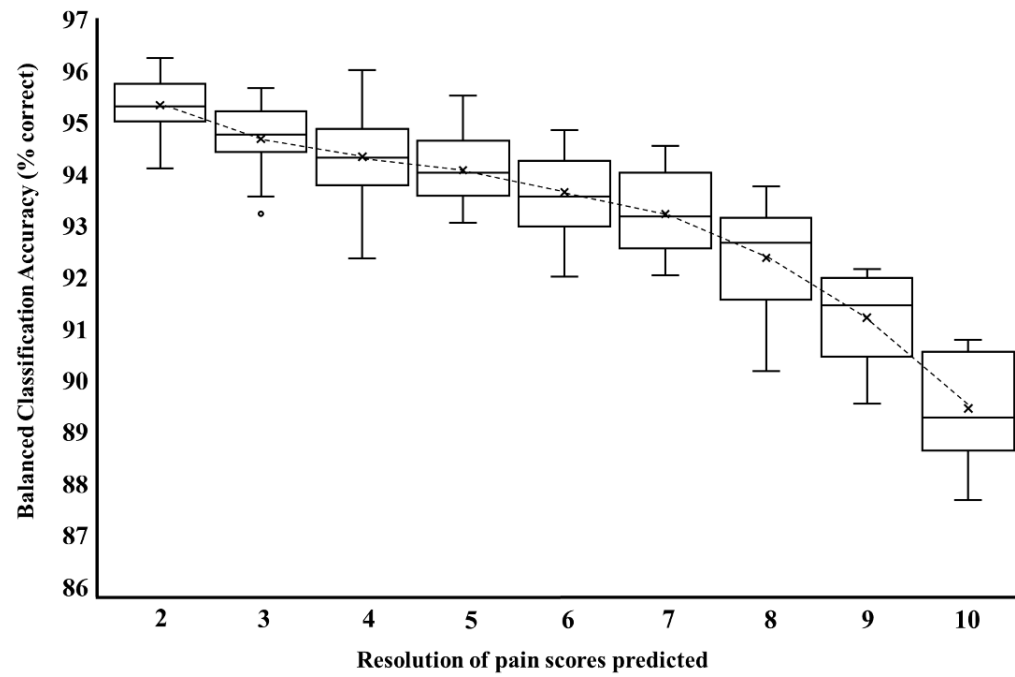
Figure 2.3: Boxplot of balanced classification accuracies (BCA) at different resolutions of the continuous pain score for classification across subjects. The X-axis depicts the results starting from 2-way classification (rest vs pain), up to and including those for 10-way classification (1-10 pain score). The low drop in BCA between 2-way and 10-way classification paradigms (6%) demonstrates the scalability of the classifier.
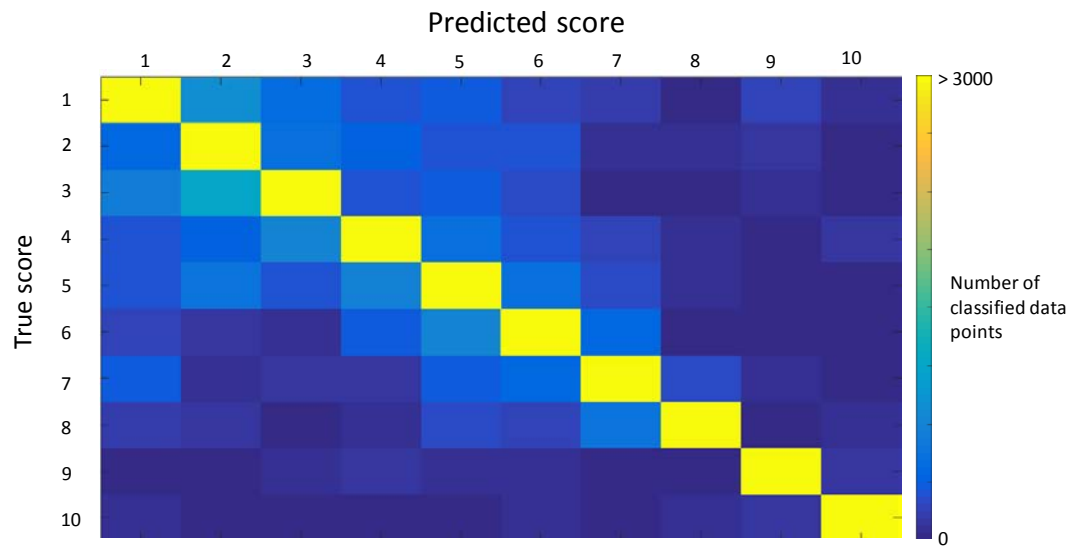
Figure 2.4: Confusion matrix for 10-way quantification using a classifier trained on 24 subjects and tested on the remaining subject. The color bar indicates the number of classified data points for the test subject.
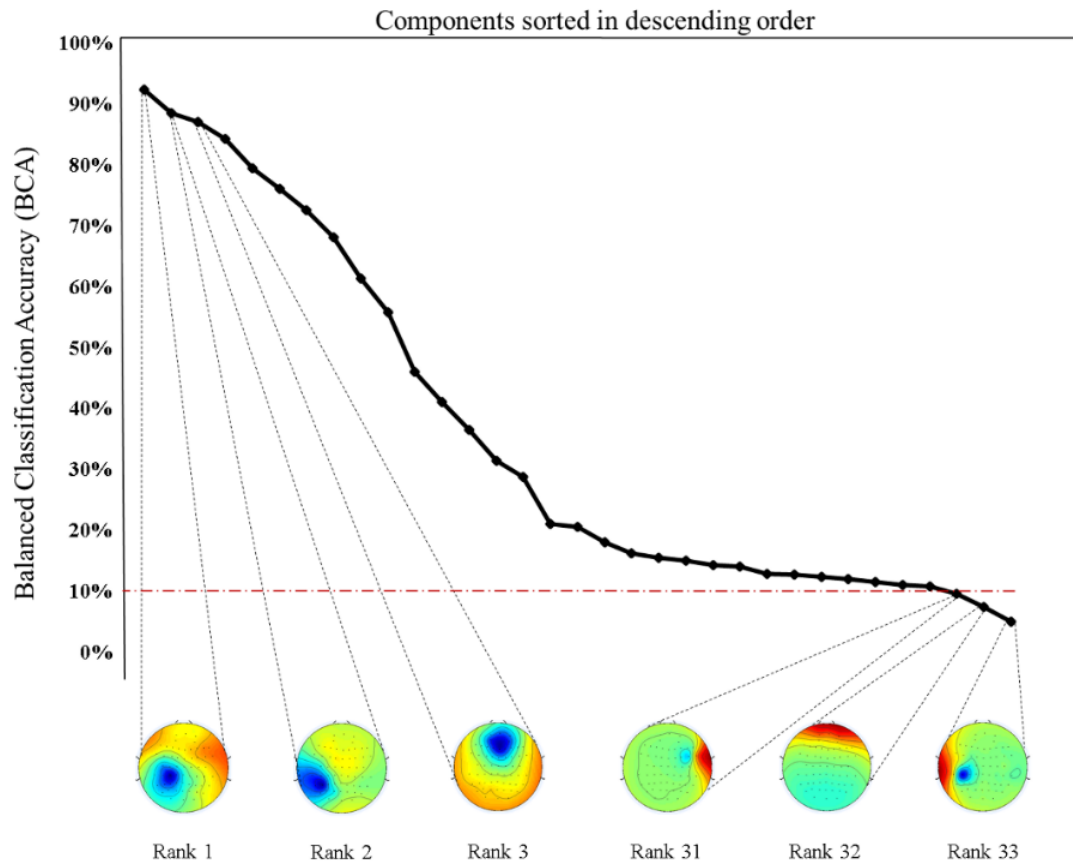
Figure 2.5: Balanced Classification Accuracy (BCA) of each component when tested on a subject, sorted in descending order of performance. The scalp maps of the best and worst three ICs are presented. The red line indicates chance BCA for the classifier. Scalp topographies corresponding to muscle activity or EEG artifacts consistently perform the worst during classification
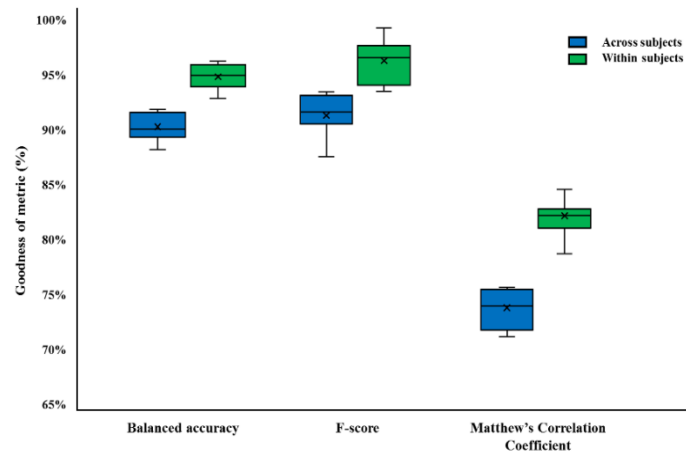
Figure 2.6: Calculating goodness of performance metrics for assessing classifier performance when trained and tested both within and across 25 subjects. All metrics are higher for the intra-subject case compared to the inter-subject case in the presence of class imbalance, as expected.
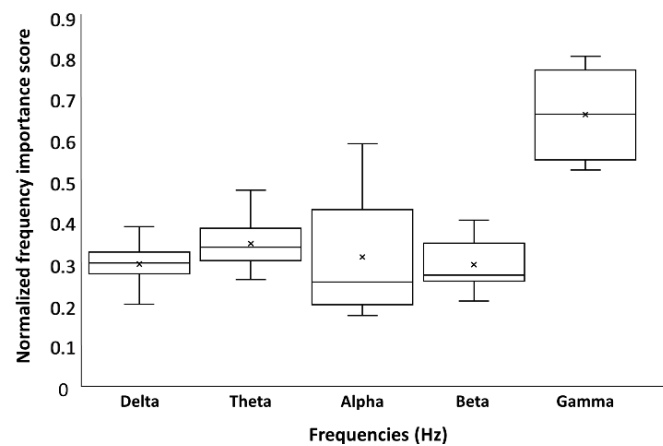
Figure 2.7: Box plot of normalized frequency contribution scores across all subjects for 10-way classification. The gamma band is seen to be consistently important across all subjects, while the alpha band shows the most variability in importance.
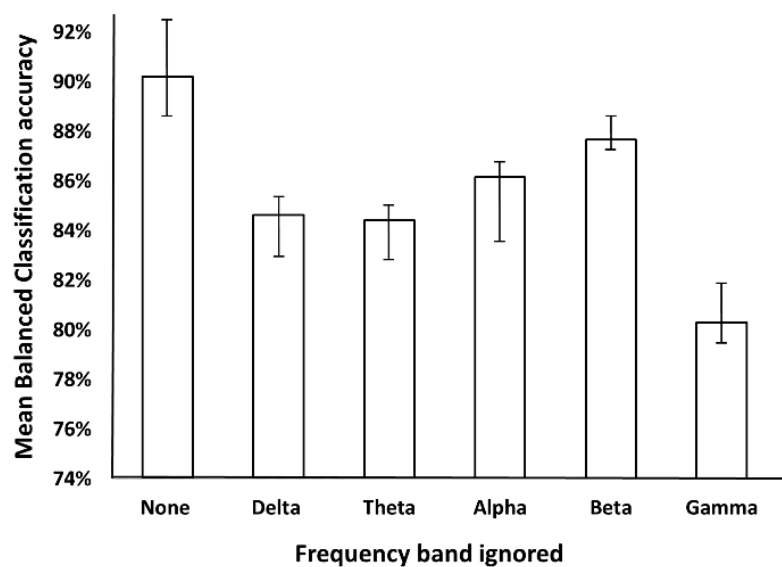
Figure 2.8: Effect of ignoring each frequency band in turn on training classification accuracy. Largest accuracy drop is seen when the gamma band is ignored, suggesting it is the most important
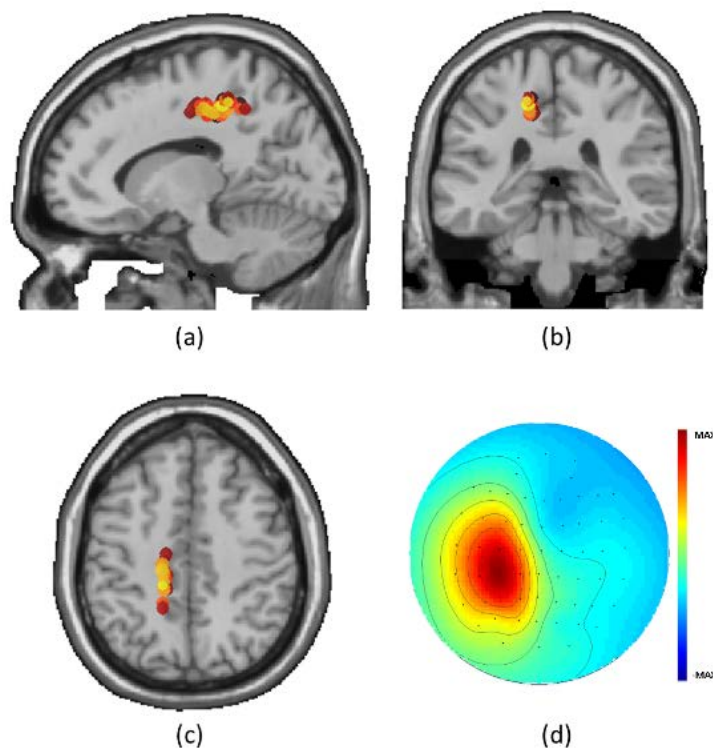
Figure 2.9: Current density source imaging of the most consistent principal component for the shared temporal pain signature. (a) Sagittal, (b) coronal, (c) axial views of source maps, and (d) scalp topography of the EEG representing the first principal component, which explains 73% of the data. The dPCC and the pMCC are the most active sources.
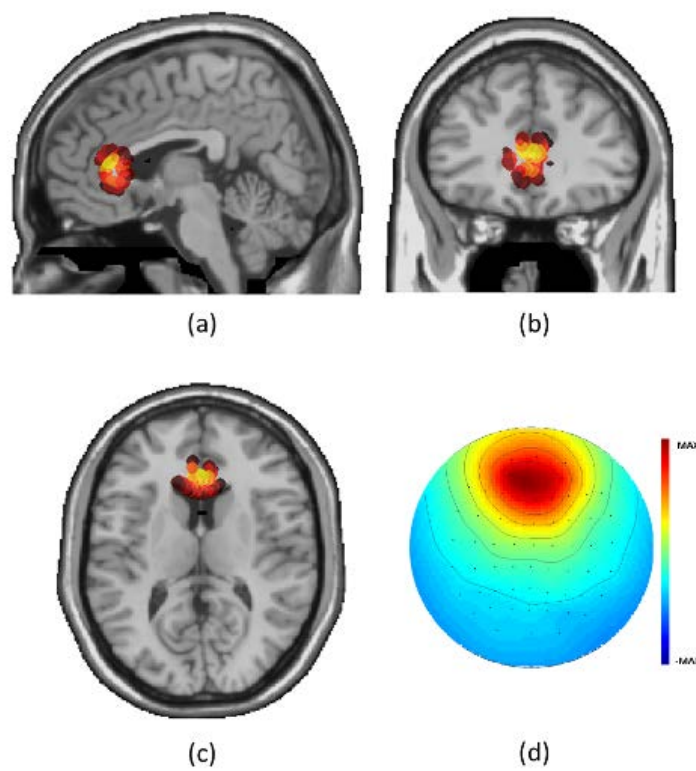
Figure 2.10: Current density source imaging of the most consistent principal component for the local temporal pain signatures. (a) Sagittal, (b) coronal, (c) axial views of source maps, and (d) scalp topography of the EEG representing the first principal component, which explains 62% of the data. The ACC is the most consistent source of activation.
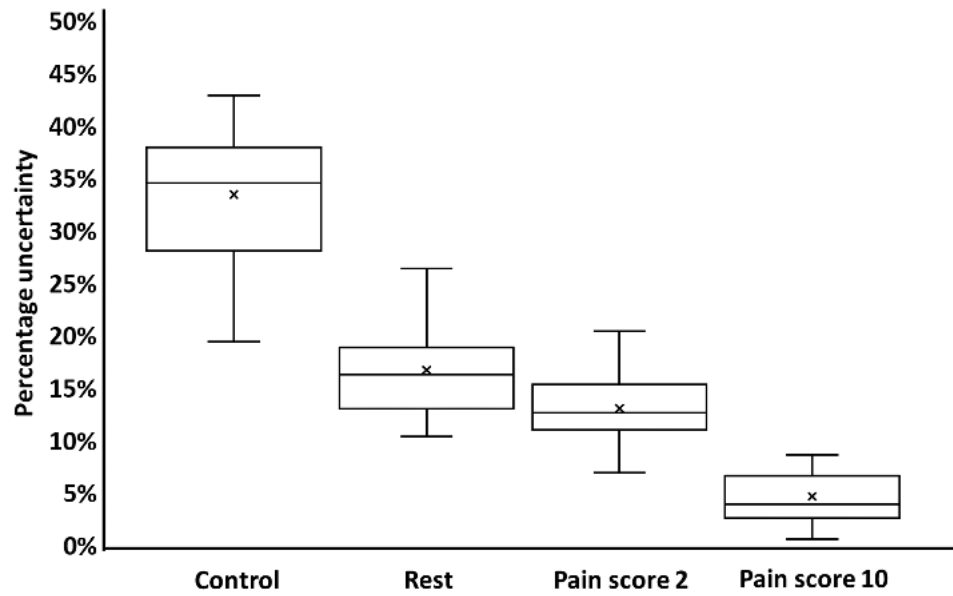
Figure 2.11: Box plot of classifier uncertainty for rest, control, and two pain scores across subjects. The uncertainty is a probability measure of how unsure the algorithm is of classifying a test point into the correct class. Greater uncertainty while classifying points belonging to the control condition shows that attention to the task plays a role during pain prediction, but low uncertainty while classifying pain scores shows that attention is not the dominant factor

Chapter 3

Deriving a Pain-Specific Neurological Signature

3.1 Introduction

Over the last four decades using a variety of neuroimaging methods, a number of regions of the brain were observed to be consistently active during pain [8]. This suggests that there is a loose spatial organization of these regions in the brain collectively called the pain neuromatrix, where the nociceptive stimulus induces a unique temporal pattern of activity leading to the feeling of pain [93]. However, due to the non-specificity of these regions, significant spatial overlap has been observed between pain and other conditions [83], [89], [94], [95]. Task-related attention was shown to be the most important confound [12] to establishing a unique pain signature due to the extent of spatial similarity between these two conditions.

Historically, establishing the presence of pain (and quantifying it) has been mainly done using functional magnetic resonance imaging (fMRI) by targeting those regions known to constitute the pain neuromatrix [6], [27], [76]. While this approach has yielded good results, the specificity of these regions to pain detection had not

been investigated. In the recent work [24], it was shown that selecting the most temporally consistent regions of activation using ICA yielded good accuracies for quantification given that there was pain. However, the ability to differentiate pain from non-painful task-related activity using information from a single source in the brain produced a significant number of false alarms, frequently indicating pain when the subject was actively engaged in a non-painful task. Several studies [8], [12] using fMRI have suggested that the regions involved in pain processing are not specific to pain, and in fact overlap with many areas that are engaged in processing and responding to task-related stimuli. This evidence suggests that a temporal investigation into the interactions between spatially distinct regions found to be active in pain processing would provide more fruitful results for pain detection.

In recent years, research using whole-brain imaging methods has led to important shifts in understanding the neurophysiology of information processing. While previous research had investigated region-specific brain behavior with the assumption that brain behavior is localized and processed in a feed-forward manner [96], an alternative hypothesis for information processing in the brain has been suggested, considering the observation that spatially distinct regions of the brain in fMRI were seen to show temporal co-activation, lasting 12-15 of seconds at a time before rapidly shifting to another spatial configuration [17], [97]. These distributed networks, called *functional networks* were observed to be active both during rest and

56

during cognitive load [98], [99]. Functional networks found during resting state, commonly called resting state networks (RSN), are hypothesized to prepare the brain for optimal processing of an impending stimulus [98], [100]–[102]. In studies analyzing their behavior [37], [103]–[105], researchers were also able to capture differences in network activity between pain and other tasks for pain detection. However, there are conflicting results about the relationship between the patterns of activation/deactivation of these networks in response to a stimulus. This could be attributed to the intuition that preparing and responding to rapid changes in the environment would suggest re-organization of the relevant RSNs in a time scale much faster than the temporal resolution of fMRI. A useful alternative would be to study these interactions in electroencephalography (EEG) using the well-established formulation of microstates.

Microstates are defined as global patterns of quasi-stable EEG scalp potential topographies that dynamically vary over time in an organized manner, lasting for 80-120 milliseconds [106], [107]. Unlike some other techniques of characterizing EEG activity, microstate analysis considers the brain activity from all electrodes to create a global representation of a functional state. Various studies have successfully correlated the spatio-temporal activity of the microstates at the millisecond level to that of much slower fluctuations in the RSNs, attributing neuroscientific significance to microstate behavior [108]. Specifically, [109] found that the four principal

microstates (A, B, C, D) estimated in previous research were associated with phonological processing (microstate A), visual imagery (microstate B), emotional salience (microstate C), and attention reorientation (microstate D). Other studies have captured more complex interactions between microstates by analyzing fractal properties in microstate time sequences during rest [110]. In addition, studies have looked at differences in microstate parameters (duration, occurrence, and coverage) during task performance compared to resting state [111], [112], observing consistent activity of the four canonical microstates. However, the functional significance of microstates has not been explored in detail, especially during pain. The objective of this study was to establish the presence of microstate behavior unique to pain that has not been seen before, and that can differentiate between thermal pain and other conditions known to evoke similar responses, such as task-related attention and non-painful heat. This unique temporal behavior can thus be used in detecting, characterizing and quantifying pain.

In this study, it was hypothesized that either microstate C or D (or both) would occur significantly more often during non-rest conditions compared to resting state due to their association with salience and attention-related functional networks. It was also hypothesized that the duration of any combination of microstates C and D would be significantly different in non-rest conditions compared to resting state. However, due to the coarseness of the conventional parameters used for microstate

analysis (duration, occurrence and coverage), it was hypothesized that they would not show any significant differences in distribution while comparing pain against non-pain conditions (non-painful heat and task). Hence, multifractal parameters extracted using the random walk formulation of the microstate time series proposed in [110] were used to investigate finer temporal differences between the four conditions. It was hypothesized that one or more of these parameters would differ significantly between pain and non-pain conditions, due to their ability to capture more complex interactions in the microstate time series.

3.2 Methods

3.2.1 Data Collection

Twenty-seven healthy subjects were recruited to participate in this study, with a median age of twenty-two years. Fourteen subjects were female. The study was approved by the Institutional Review Board at the University of Minnesota and conducted in conformity with the Declaration of Helsinki. All subjects were screened so that none of them had any history of mental illness, and they reported low levels of anxiety prior to the experiment. All subjects provided written informed consent prior to participating in this study.

3.2.2 Nociceptive Calibration

For each subject, EEG data were collected across multiple trials for four conditions: pain, rest, task, and non-painful heat. Thermal pain was administered to the subject using a thermal stimulator (ATS, Medoc, Ramat Yishai, Israel) which had a safety limit of 50°C. A custom-built analog pain-rating device was used to record their continuous pain score during the heat/pain trials as detailed in [24]. All thermal stimuli were delivered to the underside of the right forearm. To measure each subject's sensitivity to small changes in temperature, the heat delivered by the thermode was increased at a rate of 0.1°C/second starting from body temperature (37°C), until the subject indicated with the press of a mouse button that he/she felt a change in temperature. If the button was pressed for temperatures exceeding 40°C across multiple sensitivity trials, the subject was excluded from analysis. As the range of temperatures corresponding to pain is unique, the upper (pain tolerance) and lower (pain threshold) bounds on temperatures to induce pain were determined for each subject in the manner detailed in [24]. All pain trials were administered between temperatures of 1°C above median pain threshold and 0.5°C below median pain tolerance, while heat trials were administered between 38°C and 1°C below median pain threshold temperature.

3.2.3 N-back Task

For gathering attention-related EEG data, each subject was required to perform the continuous N-back task [113]. In short, the subject viewed a sequence of letters displayed in a serial fashion on a screen. This task required the subject to answer whether the current letter displayed on screen was the same as the one seen N turns ago. The sequence of letters was randomly generated using a random sample of four English alphabets. The next letter was displayed only after the subject provided a match/non-match response using two keys on the keyboard. The continuous N-back executive memory task was chosen because of the demand placed on the central executive resources of the brain, which have been shown to have significant spatial overlap with regions of the brain implicated in processing pain [114]. In addition, the complexity control afforded by this task by modulating the variable N ensured consistency in the quality of EEG data collected across subjects.

3.2.4 Experimental Procedure

For each subject, the experiment was divided into two sessions, both conducted on the same day. The first was a one-hour training session conducted between 10 am and 2 pm, while the second EEG recording session lasting two hours was conducted between 4 pm and 8 pm. All subjects reported low levels of mental exhaustion prior to each session. During the training session, each subject learned to differentiate

Experimental Protocol
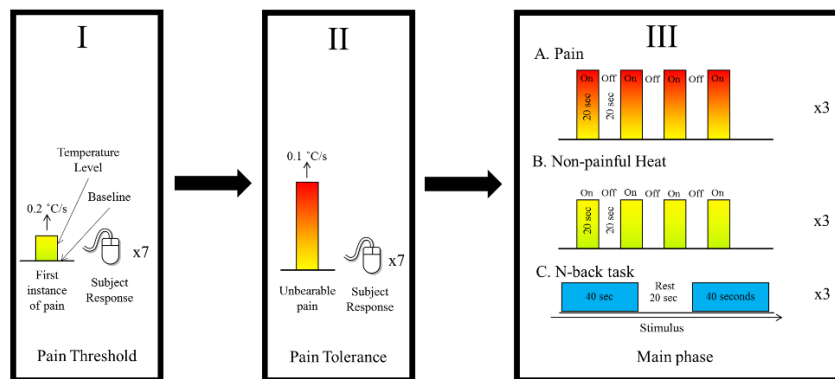
Figure 3.1: Experimental protocol. (I) Lower (pain threshold) and (II) upper (pain tolerance) bounds on pain are measured for each subject before administering separate trials. (III) Illustration of the time course of the stimulus for each condition

between the feelings of non-painful heat and pain, practiced identifying their pain threshold and tolerance using the relative pain scale, and practiced the N-back task until an appropriate level of difficulty was achieved. The complexity of the task (*N*) was adjusted based on their reaction time and the F-score calculated based on responses obtained from one minute of task performance. Unique letter and trial sequences were created for each subject, with conditions to avoid long strings of identical letters.

The framework for the second session is shown in Figure 3.1. Pain threshold and tolerance trials were conducted as detailed in [24]. Pain and heat conditions were administered for 20 seconds each, while each set of continuous task trials lasted 40

seconds. The temperature administered for heat/pain was constant to minimize the effects of salience. Four pain trials, four heat trials, and two sets of task trials were administered in each sequence. Two sequences with a break in between for recovery were administered for each subject. In a given sequence, all trials were randomly shuffled to prevent bias.

For all heat/pain trials, the subject rated the intensity of the stimulus continuously starting 1.5 seconds after stimulus administration, to minimize the effect of salience on the rating. In addition to these trials, resting state data were also collected, where each subject passively observed a silent video clip of an aquarium for six minutes to constrain their streams of consciousness.

3.2.5 EEG Data Recording and Pre-Processing

For all trials, EEG data were recorded at 1024 Hz using a 128-channel Biosemi Active-two system. The offset was kept below 20 mV for all conditions. The continuous rating was collected at a rate of 192 Hz. During the analysis of each set of pain/heat trials, the first 1.5 seconds of EEG data were removed to minimize the effects of salience due to the drastic change in stimulus intensity. The first trial for each set of task trials was also removed to eliminate effects of large hand/eye movements. Subsequent analysis was performed in MATLAB and R using native toolboxes and those developed by EEGLAB [60].

The EEG signals were down sampled to 256 Hz and band-pass filtered between 1 Hz and 100 Hz to remove any muscle contraction artifacts and voltage drifts. Line noise at 60 Hz was filtered out using a notch filter. EEG data were then re-referenced to average reference. Artifact subspace reconstruction [61] was used to clean eye blinks and reject bad time windows. This method works using sliding window principal component analysis, which estimates the signal values of high-variance components exceeding a threshold relative to a covariance matrix calculated using one minute of empirically determined clean data. Ocular and muscle artifacts in EEG were removed using blind source separation [115]. The data across all trials for each condition were then averaged. For the task condition, 20 seconds of the data were collected by selecting the middle of each set of trials, as the data were reasonably free of artifacts. The continuous EEG data corresponding to each condition were then segmented into 2 second epochs [116] for further analysis of microstate parameters. On average, 10 epochs per condition were obtained. Segmentation of the data was not performed for computation of the fractal parameters as the entire trial is used.

3.2.6 Computation of Microstate Clusters

As outlined in [117], global microstate clusters were computed using the EEGLAB implementation in MATLAB. The global brain activity can be described

by the global field power (GFP), which is the standard deviation of scalp electrical activity across the entire electrode array:

$$GFP = \sqrt{\frac{(\sum_{i=1}^{k}(V_i(t) - V_{mean}(t))^2)}{K}} \qquad (3.1)$$

The GFP represents the instantaneous strength of the electric field over the brain. The local peaks in the GFP represent instants of strong, relatively stable electric field, and highest signal-to-noise ratio (SNR) in the scalp topography. These can be considered as discrete states of the EEG, and the temporal evolution of the brain signal can be considered as a series of such states. [118]

For each subject, the microstates were computed by clustering the EEG topographical maps obtained at time points of global field power peaks in the preprocessed EEG sensor data. This modified combination of k-means and agglomerative clustering algorithms takes into consideration the high temporal correlation between successive spatial maps in EEG. This clustering algorithm was used to obtain microstate candidate maps that maximally explain the variance in the EEG map topographies. The mean microstate classes were then computed across all subjects and conditions, and are sorted according to the four canonical microstates as detailed in [119]. Each candidate scalp map and the associated EEG data point was then assigned to a cluster using a criterion called the Global Explained Variance (GEV), which is expressed as:

$$GEV = \frac{\sum_{t=1}^{t_{max}} \left(GFP(t) \cdot C_{T_t} \right)^2}{\sum_{t=1}^{t_{max}} GFP^2(t)} \tag{3.2}$$

Intuitively, this can be understood as the measure of variance contributed by each cluster weighted by the spatial correlation between C_{T_t} between the microstate map T_t and the instantaneous scalp map. Each time point was assigned to a microstate with the highest GEV. This resulted in a temporal sequence of microstate clusters that are used for further computation.

3.2.7 Computation of Microstate Parameters

As outlined in [111], the first and last microstates per epoch were discounted because of potential discontinuities disrupting the calculation of the parameters. By analyzing the sequence of dominant clusters for each epoch, the mean duration, coverage, and occurrence of each microstate were computed. The *duration* is defined as the total time for which temporally consecutive scalp maps are assigned to the same microstate cluster. The *occurrence* is defined as the mean number of distinct microstates in a given cluster occurring in a 1-second time window. The *coverage* is defined as the percentage of time covered by a given microstate. For a given microstate in each epoch, the mean duration, coverage and occurrence are computed for all trials, and subjects.

3.2.8   Multifractal Analysis of Microstate Time Series

In recent works [110], [120], [121], the autocorrelation of the microstate sequence for resting state data was shown to decay slower than the exponential distribution, implying that the sequence has long-range dependencies (LRD) and its geometrical analogue, self-similarity. Such a slow decaying autocorrelation function can be quantified using the rescaled-range measure, which is calculated by rescaling the range of the data for various time scales and dividing by the standard deviation. This measure, originally devised by H.E. Hurst, is called the Hurst exponent. A value of 0.5 indicates a memoryless process, analogous to a coin toss, while a slope greater than 0.5 indicates LRD.

Wavelets are an efficient way of combining analyses over several scales and provide a more natural analysis of scale-dependent properties in a complex system. Self-similarity and LRD are characteristic of a special set of stochastic processes that follow the power law for real exponents q, which states that such a stochastic process $X(t)$ that is scaled by a factor a is directly proportional to the unscaled process by a constant factor:

$$\mathbb{E}|X(at)|^q = |a|^{\zeta(q)}\mathbb{E}|X(t)|^q \qquad (3.3)$$

In [110], the scaling spectrum ζ was estimated for $q \in [-5, 5]$ using equation (2.3) and multifractal analysis using wavelet leaders. The first derivative of
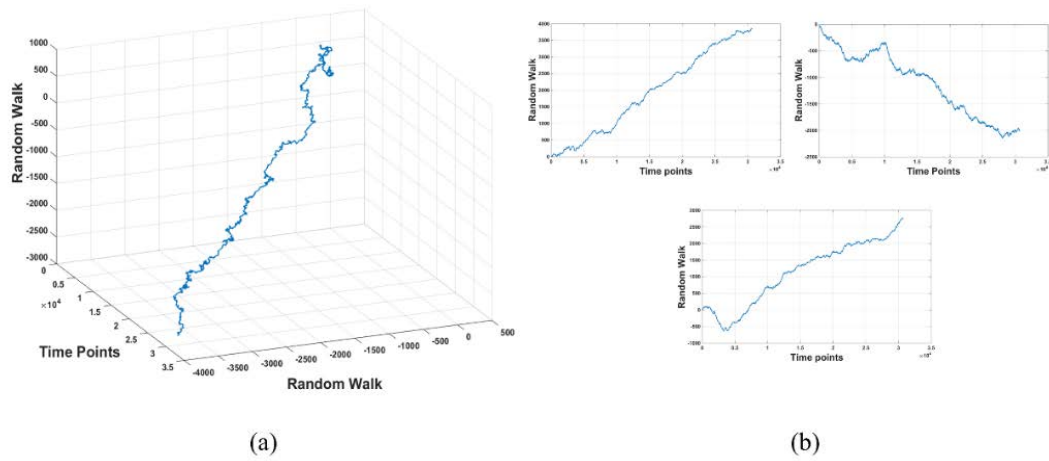
Figure 3.2: (a) Illustration of a 2-dimensional random walk constructed from four principal microstate classes, treated as outcomes of unknown, independent neurophysiological processes over time in the brain. Diagonal directions in 2D space, with time on the x-axis are treated as steps to successive microstate outcomes (A, B, C, D). (b) Decomposition of the 2D random walk into its three constituent bipartite projections of 1D random walks. For each 1D graph, axial directions in 1D space (up/down directions), with time as the x-axis are treated as steps to successive microstate outcomes belonging to a bipartite set e.g., {A, B}, {C, D}.

polynomial $\zeta(q)$ is the Hurst exponent. Using this measure, they found that the temporal patterns of the dominant microstate classes in the resting state brain were observed to be highly similar across time scales. For a given set of outcomes, time series data exhibiting this property are commonly quantified by parameters computed from fractal analysis of a random walk constructed across all the outcomes.

In this case, the time series of the dominant microstate classes was used to construct a multi-dimensional random walk with outcomes as microstate classes A, B, C, and D. The underlying neurophysiological processes governing the dominance

of a microstate for a time point are undetermined. Hence, a random walk is a good model to study the temporal dynamics of microstate behavior. This is similar to approximations seen in other fields [122], [123].

As detailed in [110], the multidimensional walk is then broken down into bipartite sets of three one-dimensional random walks for faster computation as shown in Figure 3.2 Each 1-D random walk is then analyzed across several scales using the wavelet p-leaders approach [124] by computing the scaling spectrum to illustrate the scale-dependent behavior of the data. This behavior is quantified by computing the log cumulants describing the long-range dependence properties of the temporal microstate series. The first cumulant (C_1), commonly called the Hurst exponent, is the most well studied in microstate analysis, and it quantifies the long-range behavior of the microstate time series. Geometrically, it is the slope of the curve in the scaling spectrum ζ, which is an indication of monofractality. Higher order cumulants that are computed (C_2, C_3) describe deviations from mono-fractal behavior such as curvature and asymmetry (respectively) of the curve in the scaling spectrum.

3.2.9 Statistical Analysis of Microstate Parameters

The objectives of statistical analysis on microstate parameters were to look for significant effects of condition (pain, rest, non-painful heat and task) on one or more

69

of the microstate parameters, for each microstate. The normality of the distribution of each parameter across all subjects and conditions was tested using a Lilliefors test. As the null hypothesis of normality was not rejected, a 4 x 4 repeated measures analysis of variance (rmANOVA) was carried out for each parameter, with one factor as the condition, and the other factor as the microstate class. To minimize the risk of type I errors due to dependence across the microstate parameters, Bonferroni corrected significant results are used to conduct post-hoc paired t-tests for all condition pairs and microstates to look for significant interactions. For significant interactions between conditions and microstates in the rmANOVA, six paired t-tests were conducted per microstate (rest vs heat, rest vs pain, rest vs task, heat vs pain, heat vs task, pain vs task), with p-values corrected for type I error using Bonferroni correction.

3.2.10 Statistical Analysis of Multifractal Parameters

The objective of statistical analysis on the mean distribution of log-cumulants for different conditions (rest, non-painful heat, pain, and task) was to look for evidence supporting the hypothesis that one or more of the log cumulants calculated for each condition across subjects are significantly different across conditions. The distribution of log-cumulants calculated across subjects was tested for multivariate normality using the Shapiro-Wilk test. Box's test of equality was used to verify

equality of the covariance matrices across the log-cumulants. As neither test rejected the null hypothesis, a multivariate ANOVA (MANOVA) was used to look for significant differences across all conditions and subjects for the three log cumulants. Subsequent post-hoc univariate ANOVAs and Tukey-Kramer tests were conducted to look for pairwise significant differences between conditions and each log cumulant.

3.3 Results

The four microstates computed across all conditions are shown in Figure 3.3, and the microstates computed across all subjects for each condition are shown in Figure 3.4. Four microstate classes were used for the analysis as increasing the number of clusters did not yield a significant increase in the percentage of explained variance in the data, as seen in Figure 3.5. Across all participants and conditions, four microstate classes explained 79.64% ± 3.6% of all time points in the data.

3.3.1 Statistical Analysis on Microstate Parameters

Across all conditions, the assumption of sphericity was violated across all microstate parameters and classes during the rmANOVA. Hence, all p-values are reported with Greenhouse-Geisser correction. All p-values during the rmANOVA are also corrected for type I error using Bonferroni correction since duration,

occurrence and coverage are dependent measures. As shown in Table 3.1, the rmANOVA showed significant interactions between condition and microstate for microstate duration and occurrence, but not for microstate coverage.

3.3.2   Interactions for Microstate Durations

The 4x4 rmANOVA showed significant results for condition x microstate, with $F_{(13.18, 33.657)} = 2.984$, $p = 0.003$. Post-hoc paired t-tests revealed significant differences in the mean durations of microstate D and microstate C between pairs of conditions, as shown in Figure 2.6. The duration of microstate D was significantly higher ($p < 0.05$) during rest (82.8 ± 3.9 ms) compared to non-rest conditions. However, the duration of microstate C was significantly higher ($p < 0.05$) during rest (77.78 ± 3.6 ms) compared to heat and task conditions, but not during pain (72.6 ± 4.2 ms).

3.3.3   Interactions for Microstate Occurrences

The 4x4 rmANOVA showed significant results for condition x microstate, with $F_{(14.008, 36.069)} = 2.65$, $p = 0.033$. Post-hoc paired t-tests revealed significant effects of microstate D and microstate C on pairs of conditions, as shown in Figure 3.7. During rest, it was observed that microstate D occurred significantly ($p < 0.05$) less often (4.1 ± 0.36 microstates/second) during rest compared to all non-rest

conditions. On average, microstate C occurred more often in non-rest conditions compared to rest. However, it occurred significantly ($p < 0.05$) less often during rest (3.99 ± 1.39 microstates/second) compared to the pain and task conditions, but not during heat (4.46 ± 0.5 microstates/second).

3.3.4   Statistical Analysis on Multifractal Parameters

The MANOVA test, shown in Table 3.2, showed significant results for condition x log-cumulant, with $F_{(26, 182)} = 21.474$, $p < 0.001$. Thus, post-hoc univariate ANOVAs were conducted for each log-cumulant, with the reported p-values corrected for type 1 errors using Bonferroni correction, with the results shown in Table 3.3. The ANOVA for the first log-cumulant (C_1) reported significant differences in mean value across conditions, with $F_{(3, 36)} = 142.28$, $p < 0.001$. Similarly, the ANOVA for the second log-cumulant (C_2) was also significant across conditions, with $F_{(3, 36)} = 3.873$, $p < 0.05$. However, the third log-cumulant (C_3) failed to show any significant difference in mean across conditions. Subsequent post-hoc Tukey-Kramer tests for each significant log-cumulant expound on these differences across each pair of conditions.

As illustrated in Figure 3.7 for C_1, it was seen that significant differences in mean value were found for pain vs non-pain, and rest vs non-rest comparisons. Specifically, the value of C_1 was significantly ($p < 0.001$) higher for pain ($0.63 \pm$

0.014) compared to non-pain conditions, while it was significantly (p<0.001) lower compared to rest (0.772 ± 0.032). Consequently, the mean and standard deviation of C_1 for rest was significantly higher (p < 0.001) compared to non-rest conditions. However, the mean of the distribution of the C_2 across subjects was significant (p < 0.05) only for pain (-0.02 ± 0.011) vs task (-0.005 ± 0.017), showing that pain shows significantly higher deviation from monofractality compared to the task condition.

3.4 Discussion

In this study, we investigated deviations in the temporal behavior of microstates in 27 healthy subjects experiencing thermal pain, compared to resting state. In addition, we also researched effective ways of modelling the microstate time series to differentiate pain from other conditions known to recruit spatially overlapping mental resources for stimulus processing and response (non-painful heat and task-related attention). We looked into characterizing microstate behavior for somatosensory stimuli (pain or non-painful heat) with continuous engagement of the visual and attentional components of the brain. We hypothesized that microstates C or D (or both) would have significantly different distributions of parameters computed during rest, compared to non-rest conditions. However, we hypothesized that the distributions of these parameters would not be significantly different for pain compared to non-painful conditions that required attention to the stimulus. We also

found that quantifying subtler interactions between the microstate classes across various time scales using fractal parameters computed from the temporal microstate sequence yielded significant differences between pain and non-painful conditions, including rest.

3.4.1 Behavior of Microstate D across All Conditions

Microstate D was shown in [109] to be positively correlated with activity in the right-lateralized dorsal and ventral areas of the frontal and parietal cortex in the brain. The resultant activity of the ventral fronto-parietal network (FPAN) was shown in [125] to be strongly modulated by reflexive aspects of attention and detection of deviant stimuli, while the dorsal FPAN was active during mental reorientation and switching of attention between different tasks [126]. While showing a mean increased activation in the bilateral FPAN using fMRI, the stability of these networks was not investigated in previous studies [127]. We found that the duration of microstate D was significantly longer during resting state, compared to all non-rest conditions, while the occurrence of this microstate was significantly higher for non-resting conditions compared to resting state. We did not observe any significant differences in the distribution of parameters of this microstate during pairwise comparisons between non-rest conditions in post-hoc t-tests. In addition, we did not

find any significant distributions in microstate coverage for any condition, likely due to the contrasting trends of duration and occurrence.

This behavior is interesting because it implies that the activity of this microstate is more chaotic than previously observed [111], [112]. In [112], both duration and occurrence were significantly lower during task-related attention compared to rest whereas in [111], both the duration and the occurrence of this microstate were higher during resting state. This could be attributed to the nature of the tasks administered in each case. In [112], the task was a serial-sevens subtraction task where the subject stared at a fixed point on the screen and mentally subtracted in steps of seven from a given number whereas in [111], the task was to view an image/text on screen for a fixed duration (50 seconds) before being mentally engaged. Our attention-based task paradigm combines the functional aspects of these two studies.

Specifically, for non-rest conditions, the N-back task continuously engaged the attention of the subject while the letters changed often, forcing continuous reevaluation and mental recall. During both painful and non-painful heat trials, the thermal stimulus was maintained at constant temperatures while the continuous rating showed changes in perception of stimulus intensity throughout each trial. This would suggest fluctuations in the level of attention paid to the stimulus, a phenomenon attributed to the complex interactions between the salience network (stimulus recognition), the dorsal attention network (stimulus rating), and the default

mode network (baseline resting) as detailed in [128]. The lack of significant differences in parameter distributions between the non-rest conditions for this microstate is supported by [111] for three different task-related attention paradigms.

These results taken together would suggest more consistent activity of microstate D as observed by the significantly higher number of occurrences during non-rest conditions, but the stability (duration) of this microstate during task, pain, and non-painful heat conditions would need to be investigated further.

3.4.2 Behavior of Microstate C across All Conditions

Microstate C was shown in [109] to be correlated with activity in the pregenual anterior cingulate cortex (pACC) and the right anterior insula (AI), among others. Taken together, these regions were hypothesized to form the fronto-insular network, a core part of the salience network. The main function of this network is to switch between the activity of the default mode network (DMN) and the central-executive network (CEN) [129]. In other words, its role is to detect deviant stimuli, engage relevant brain areas responsible for attentional, working memory, higher order cognitive processes, and to engage/disengage the default mode network (DMN) [17]. Hence, the role of functional networks attributed to microstate C share similarities with that of functional networks attributed to microstate D.

This conclusion is supported by our results, which shows that the trends observed in duration and occurrence for microstate C are similar to microstate D for rest vs non-rest conditions, with higher duration in rest and higher occurrence in non-rest conditions. However, for microstate duration in C, significance is achieved only in heat vs rest and task vs rest conditions, while the distributions of occurrence in C are significantly different for pain vs rest and task vs rest conditions. Again, no significance is found for microstate coverage for any pairwise comparison of conditions due to differing trends in duration and occurrence.

The higher occurrence of microstate C is in agreement with previous findings that the components of the salience network (AI and the pACC) are consistently active and play different roles while switching between the CEN and DMN [17]. Specifically, the AI was hypothesized to play a more prominent role in salient stimulus detection, while the pACC is responsible for modulating responses in the sensory, motor, and association cortices. These types of responses are expected during performance of the N-back task, the experience of painful/non-painful heat, and the consequent action of rating it. However, the inconsistency of significant occurrences across all rest vs non-rest conditions may be due to its role in engaging/disengaging the DMN, a network that is active during rest. The higher duration of this microstate during rest compared to other conditions could also be explained by its association to the DMN [112]. Hence, the behavior of microstate C

in this study is another example of more chaotic activity compared to other studies for the rest vs task-related attention case [111], [112]. As outlined in [107], further studies need to be conducted into the activity of this microstate due to its seemingly inconsistent behavior across all conditions.

The behaviors of microstates C and D across all the conditions in this study showed that the use of first-order parameters like duration, occurrence and coverage were not sufficient to differentiate between non-rest conditions effectively. In addition, the inconsistent nature of significant trends established across individual microstates do not provide any new knowledge of microstate interactions. However, the similarity in trends across these microstates and the interplay between the functional networks that they represent would suggest that a formulation that considered interactions between all four microstates would be better suited to illustrate the temporal differences across one or more of these conditions. Hence, analyzing the chaotic nature of microstate sequences across several time scales proved a better alternative for identifying unique patterns, as illustrated in the next section.

3.4.3 Multifractal Analysis of Microstate Sequences

Inspired by similar formulations in [122], [123], the examination of the temporal microstate sequence across several scales as demonstrated in [110] yielded more

unique results in terms of the stability of temporal microstate behavior. Specifically, while all conditions showed long-range dependency, the means of the distributions of the Hurst (C_1) parameter across all subjects showed significant differences for both pain vs non-pain conditions, and rest vs non-rest conditions.

Scale-free dynamics (or monofractality) in the resting-state brain, as summarized in [110], [121] is generally regarded as a sign of efficient and flexible flow of information between multiple sources in the brain. Self-similarity across various time scales is an indication that the microstate time series has a clear structure in the temporal organization that is not completely random, but not completely predetermined. Self-similarity, quantified by C_1 can hence also be considered as a measure of correlation across time scales between the duration of a microstate and the corresponding functional network in the fMRI time scale. For this parameter, $0.5 < C_1 < 1$ indicates long-range dependency. To date, task-specific disturbances in resting state network dynamics have been observed only in fMRI [97], [130]. This is the first study to demonstrate significant differences in scale-free behavior for task-specific conditions in EEG.

It is interesting to note that while resting state has maximum self-similarity (mean $C_1 = 0.77$) among all conditions, there is a significantly higher correlation in the pattern of microstates across scales for pain (mean $C_1 = 0.63$) compared to task-related attention. This would imply that the sensation of pain is processed in a similar

manner for different time scales, and there is an inherent structure in the microstate sequence that is absent while processing non-painful thermal stimuli, or while performing a task. The latter phenomenon could be the nature of the task performed, as the subject completes each trial of the N-back task at the rate of 0.47 trials/second. This would indicate continuous mental reorientation, which would exhibit dissimilar patterns across several time scales within the duration of the trial, hence leading to relatively poor correlation (mean $C_1 = 0.544$) across time scales. In addition, as pointed out in [128], mind-wandering studies conducted using fMRI on subjects experiencing thermal stimuli below or near their pain threshold (when they start to feel pain) show that highly variable pre-existing brain states determine their response over time. For our study, this would induce highly dissimilar patterns of processing non-painful heat stimuli, which may be why it shows relatively poor correlation (mean $C_1 = 0.547$) across time scales. However, we administered thermal stimuli found to induce medium-to-high painful responses in each subject determined before the experiment during the training sessions, which would inherently reduce inter-trial variability.

3.5 Conclusions and Future Directions

The aim of this study was to identify a temporal pattern of microstates that is unique to pain, which can differentiate between other stimuli known to evoke similar

spatial responses. We analyzed the differences in the temporal microstate sequences extracted from experimental stimuli designed to capture attentional, somatosensory and painful correlates of stimulus processing. We found that the trends of first order microstate parameters such as microstate duration and occurrence, while significantly different for non-rest conditions compared to rest, were too chaotic to be effectively unique to pain. We found that a multifractal approach to modelling the randomness of the microstate sequences across increasing time scales illustrated the presence of microstate behavior unique to pain. This work is a step forward in effective characterization of a unique temporal pain response that can be used for pain detection and quantification.

While the results found in this study are encouraging for further pain research using microstates, the results found by analyzing microstate behavior across time scales warrant further research. Specifically, investigations into the correlation decay across time scales, and conducting source-imaging studies on the EEG data at the time scale where maximum decorrelation was observed would shed light on the behavior of individual microstates, or dependencies between them. Thus, stronger ties can be established between temporal patterns at the sub-second scale to the behavior of resultant resting state networks.
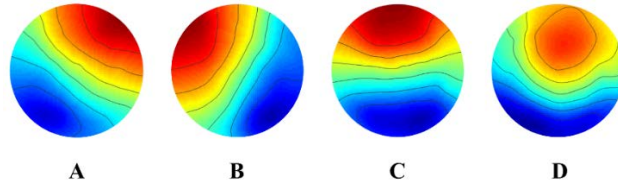
Figure 3.3: Global microstate classes, calculated across all subjects and conditions
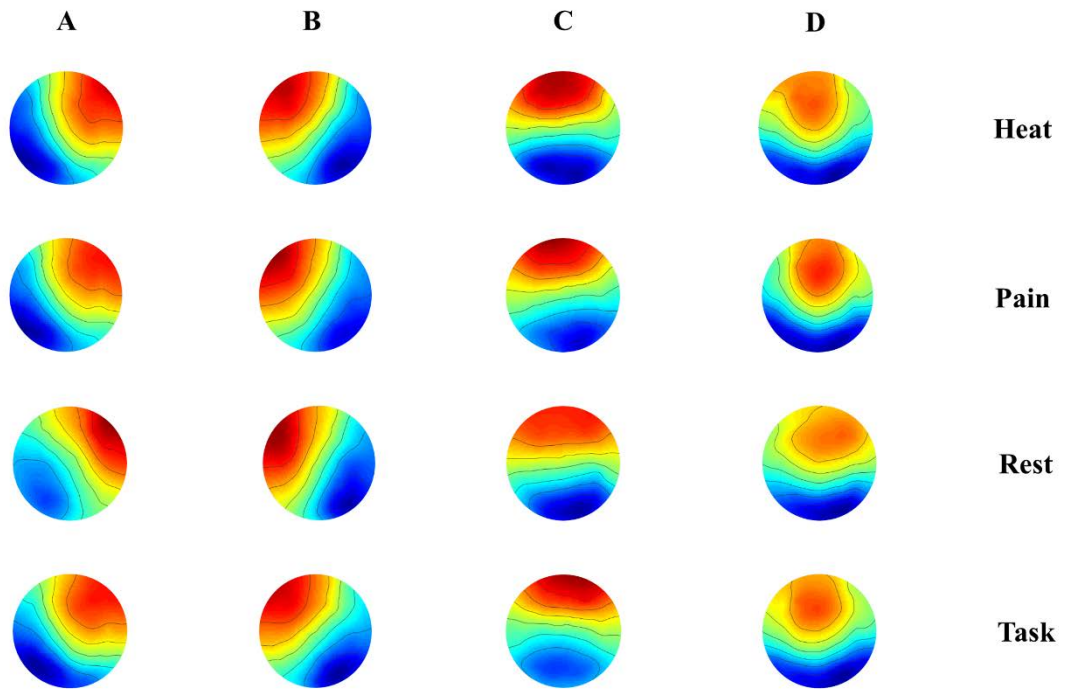
Figure 3.4: Grand mean microstate classes calculated for each condition across all subjects
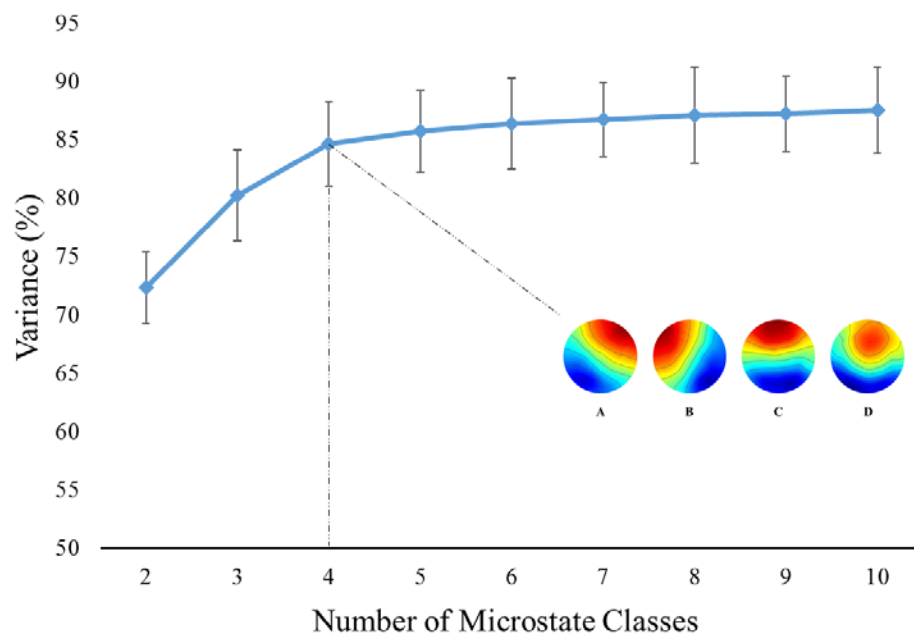
Figure 3.5: Percentage of variance explained in the data as a function of candidate microstates across all conditions. The number of clusters is increased from two to ten microstates, with the knee of the curve found using four microstates.
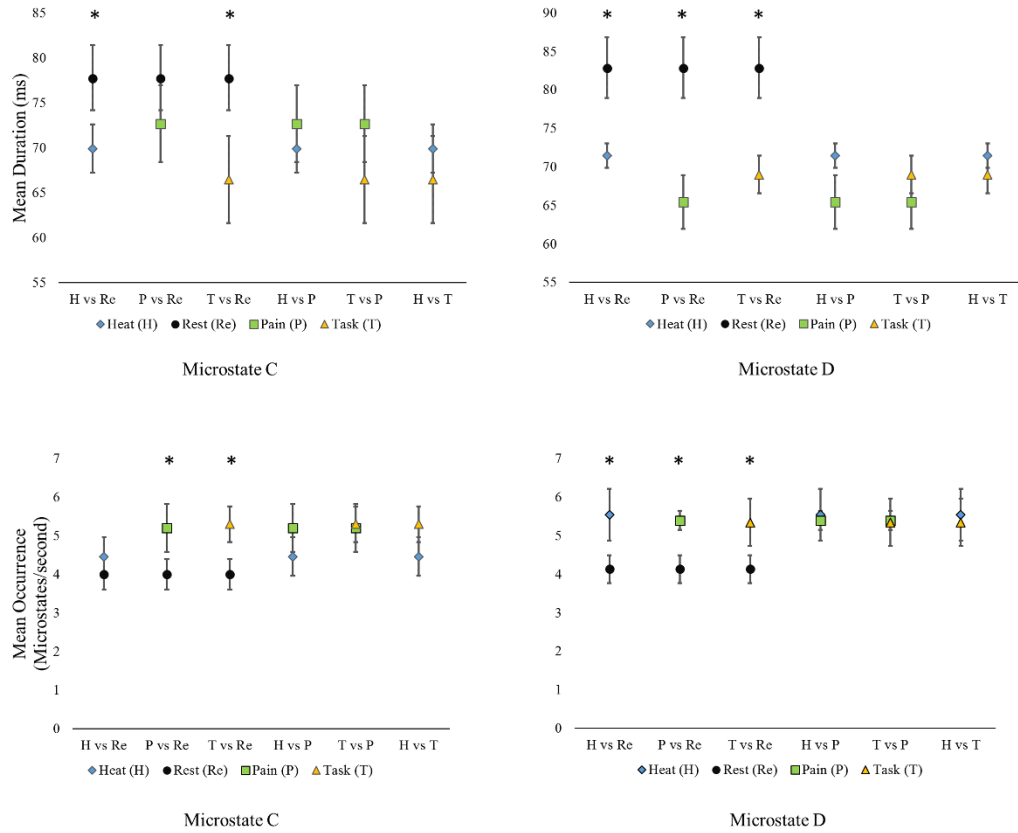
Figure 3.6: Plots illustrating the results of paired t-tests conducted across pairs of conditions (across pain, non-painful heat, task, and rest) for microstate duration and occurrence for the two significant microstates (C and D). The top row depicts the mean and standard deviation of durations across conditions. The bottom row depicts the mean and standard deviation of occurrence across all conditions
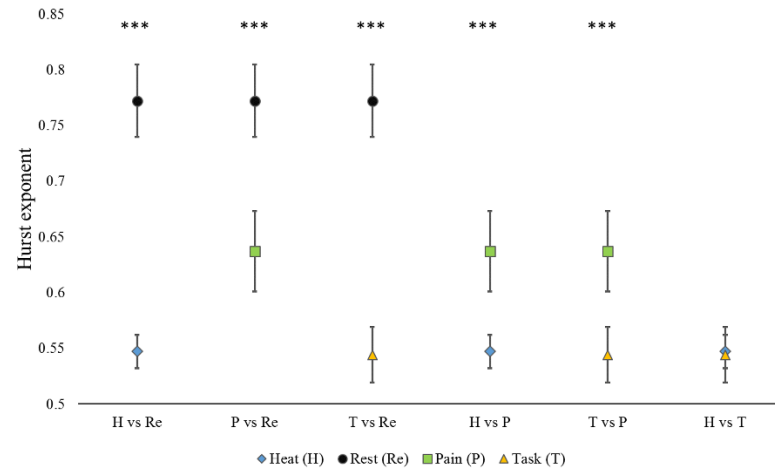
Figure 3.7: Plot illustrating the mean and the standard deviation of the Hurst parameter (C1) for pairwise comparison across pain, non-painful heat, task and rest.

Table 3.1: Results of rmANOVA on interactions between condition and microstate parameters

Parameter	df	df error	F	p value	partial η^2
Duration	13.18	33.657	2.984	0.003	0.247
Occurrence	14.008	36.069	2.65	0.033	0.231
Coverage	14.49	40.41	0.96	0.48	–

Table 3.2: Multivariate effects of log-cumulant on condition

Dependent Variable	df	df error	F	p value	Wilk's Λ	partial η^2
Log-cumulant	26	182	21.474	<0.001	0.053	0.623

Table 3.3: Univariate effects of each log-cumulant on condition

Parameter	df	df error	F	p value	Partial η^2
C1	3	36	142.23	<0.001	0.922
C2	3	36	3.873	0.011	0.244
C3	3	36	1.344	0.275	0.101

Chapter 4

A Deep Learning Approach to Objective Pain Detection

4.1 Introduction and Motivation

Most modern objective pain detection algorithms are implemented as an attempt to classify the incoming data as "pain" vs "no pain", where the "no pain" class usually does not consider more than a standard resting state condition. This has yielded many machine learning algorithms for both pain quantification and detection using fMRI and EEG data [18], [23]–[25], [27]. However, in order to be deployed as a viable tool beyond the confines of the clinic, a pain detection algorithm should focus on specificity, where it needs to ensure that different types of responses to non-painful stimuli are actually classified as "non-painful". Such a detector would aid in recording and comparing the onset, duration and intensity of painful episodes in longitudinal clinical trials designed to test the efficacy of certain pain therapies over time for patients. In such cases, a significant amount of computational resources would be dedicated to sifting through and discarding data unrelated to pain. However, the main confound to detecting pain is task-related attention, as attention to painful stimuli and attention to task-related stimuli tend to show very similar

evolutions of the spatiotemporal neurophysiological signature. Although task-related attention and nociceptive modulation of attention-related neurological responses have been studied extensively in pain research [12], only recent discoveries analyzing functional networks in both fMRI [128] and EEG [107] have revealed differences in these response patterns. Task-related attention is a very common occurrence in a real-world scenario [114], which is why an attempt needs to be made to develop a robust machine-learning algorithm that can effectively detect pain-specific responses, even in the presence of other non-painful confounds.

Even though a pain-specific spatiotemporal biomarker using EEG microstates was statistically established in chapter 3, the number of optimal microstates is dependent on the dataset, which influences the calculation of the wavelet p-leader estimates for multifractal analysis. This would require more samples proportional to an increasing number of microstates in order to obtain stable, reliable estimates of the log cumulants. For example, the reformulation presented for four microstate classes required about 4000 input samples (~18 seconds of continuous EEG data per stimulus) to produce reasonably stable estimates. Hence, this approach would not be optimal for fast pain detection when considering parallel analysis of the activity of multiple microstates. However, alternative methods that build upon and extend this conclusion can be used to build a robust and effective pain detection system.

While EEG microstate sequences are seen to encode a pain-specific temporal pattern, one spatial map configuration (such as a microstate) does not necessarily define the entire global state for a set period of time. This is validated by studies employing other clustering methods to find potential microstates, such as ICA [131], PCA [132] and successive iterations of the k-means clustering algorithms [133], which do not consistently produce the same dictionary of microstates that provide the best fit for that time point, pointing to a lack of reproducibility of the analysis. Hence, finding only the best fit of a microstate to a spatial map based on the spatial correlation coefficient may not completely define the neurophysiological behavior at that time point. An alternative would be to generalize this approach, analyzing the temporal changes in the *distribution* of the fit of a dictionary of microstates over successive instantaneous EEG spatial maps. This data-driven approach can be used to model the temporal dependencies between these spatial maps more effectively, which would present a more complete picture of brain activity specific to pain.

A potential confound to this approach is that classical EEG microstate analysis results in candidate microstates that show strong pairwise temporal correlations [134]. Such a pattern of activity would be detrimental for objective pain detection when these clusters are analyzed in the context of variation of distributed activity over time. Using spatiotemporal ICA on the GFP maxima to compute candidate microstates [131] is a more meaningful alternative to obtain microstate maps that

show temporal independence. This can be used to form the microstate dictionary for pain detection.

The more generalized approach outlined above necessitates an exploration into deep learning methods for automatic temporal feature abstraction, given the inputs as the instantaneous spatial maps, with the feature descriptors for each input as the fit of each candidate microstate cluster to the instantaneous spatial map.

4.2 Artificial neural networks

4.2.1 Introduction

Artificial neural networks (ANN) are a class of supervised machine learning algorithms for classification and regression [135]. The formulation of ANNs draws inspiration from connections formed between biological neurons in the nervous system in order to transmit signals. The base of a neuron is formed by a branching tree of dendrites, which receive signals from other neurons. These signals are then shaped in the cell body of the neuron, which are then propagated forward to other neurons through a long shaft, called an axon. The extent of learning is determined by the robustness of connections formed between successive layers of neurons for a given stimulus.
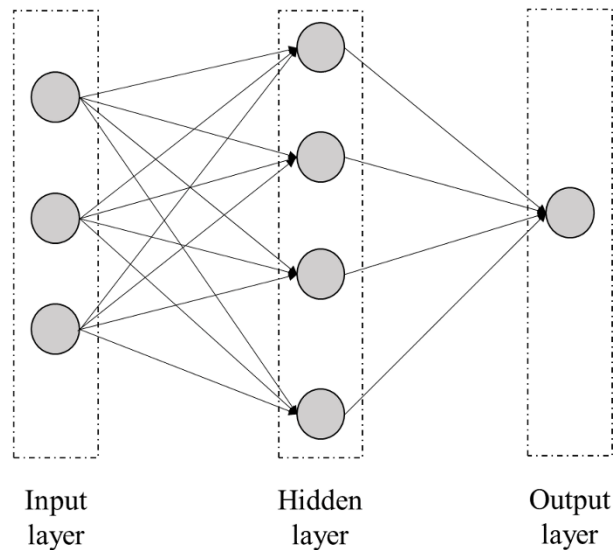
Figure 4.1: Visualization of an artificial neural network (ANN)

ANNs are designed to approximate this behavior. The basic unit of an ANN is called a node. Each node performs a computation by weighting the inputs from previous layers in a non-uniform fashion, and propagates the output to successive layers, as shown in Figure 4.1. Hence, ANNs are arranged as a feedforward graph with the first layer receiving the input, successive layers modifying the input, and the final layer producing the output. The number of layers between the input and the output layers are configurable and are called *hidden* layers due to their outputs being encapsulated. Each node in a hidden layer can be connected to a variable number of nodes in the previous layer. The number of outputs in the output layer is determined

by the number of classes in the case of classification, or it can be a single node producing a real-valued output in the case of regression.

In each node, the output is a weighted linear summation of the inputs, with the intuition that the weights computed at a given node represent a measure of importance assigned to the output of   each node from the previous layer. More formally, the computations carried out by the node i in layer j is characterized by the weight vector $w^{i,j}$ and the bias $b^{i,j}$. For inputs a^{j-1} from the previous layer, the output is given as

$$z^{i,j} = w^{i,j} . a^{j-1} + b^{i,j} \tag{4.1}$$

The output of the node is usually passed through a non-linear function $f(z)$ to produce an output $a^{i,j}$ that is passed to the next layer. Some common non-linear functions are the sigmoid function, the hyperbolic tangent function (tanh), the softmax, and the rectified linear unit (ReLU) function:

$$f(z) = \frac{1}{1 + e^z} \tag{4.2}$$

$$f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{4.3}$$

$$f(z) = max(0, z) \tag{4.4}$$

$$f(z) = \frac{e^{z_i}}{\sum_k e^{z_k}} \tag{4.5}$$

Sigmoid and tanh activation functions are bounded and differentiable, hence used in multiple hidden layers. Rectified linear units are most commonly used because they provide a good approximation of the behavior of a biological neuron. The tanh activation function is desirable when using neural networks with many hidden layers. In the softmax function, the activation at a node is normalized with the respect to the current layer, where indices i and k address all nodes in the same layer. Hence, this activation function is typically used for multi-class classification.

Traditional machine learning algorithms perform best when trained on increasingly sophisticated characterizations of the training data, called features. The efficiency of the classifier in learning the associations between these features, as well the association between the feature vector and the class label determines its degree of performance when testing on an unseen, independently collected datapoint. Thus, the limit of traditional classifier performance is conditioned on the meaningfulness of features extracted. ANNs have been shown to surpass this limit by forming abstractions of the training data within successive hidden layers and continuously updating the weights of each layers proportional to the magnitude of misclassification. The input layer for an ANN contains as many nodes as there are features for a particular datapoint. Modern deep learning problems can have more than a million features, where some initial preprocessing is performed on the data

before feeding it to the ANN to engineer features specific to the classification problem.

4.2.2 Backpropagation

In the first pass through the neural network, the weights for all nodes in each layer are initialized in a pseudo-random fashion, and the feature vector is propagated through all the layers to the output. The difference between the predicted output and the observed output is quantified by the objective function called the *loss*, which is a measure of the magnitude of misclassification. Minimizing the loss through an iterative procedure hence trains the network over time to predict the target outputs accurately. This process, called backpropagation [136], is to apply the chain rule of the derivative of the loss function in order to estimate the weights for each layer of the network.

The most commonly used loss function for neural networks is the logistic loss, or the cross-entropy loss. For m output classes, n training examples, target class y and predicted class $\hat{y}$, the cross-entropy loss is given as:

$$L = \sum_{n} \sum_{m} y_{mn} \, ln(\hat{y}_{mn}) + (1 - y_{mn}) \, ln(1 - \hat{y}_{mn}) \qquad (4.6)$$

The update rule for the weight vector **w** for a given layer is updated as:

$$\boldsymbol{w} := \boldsymbol{w} - \eta \nabla L(\boldsymbol{w}) \qquad (4.7)$$

The derivative of the loss function L is computed with respect to the weight vector $\mathbf{w}$ using the chain rule, and the learning rate η, which governs the step size. Choosing the learning rate affects the computational time and efficiency required to converge to a global minimum.

In short, the forward computation of the weight vectors is achieved using (4.1) and the activation at each node for each layer is computed using (4.2-4.5). The loss at the output node is then computed using (4.6) and the weights are then updated according to (4.7). This process is repeated until the loss function converges to a global minimum, or the upper limit on the number of iterations is reached.

Due to the extensive computational resources needed to perform each update step for all nodes and layers and for all training examples, many alternatives to gradient descent have been proposed. The most popular among these are the stochastic gradient descent (SGD) and the mini-batch gradient descent algorithms. The SGD variant chooses to optimize over a randomly selected training example at a time, with many wide jumps observed at each iteration, with the intuition that fast convergence would depend on a handful of informative training examples, hence resulting in bigger jumps towards the global minimum of the loss function. Due to large oscillations observed near the minimum using this approach, mini-batch gradient descent is a useful alternative, where backpropagation is performed over a batch of randomly selected training examples, in order to make gradient descent

smoother and to speed up computation time. These approaches, coupled with the development of memory-efficient computational hardware, has resulted in deep learning algorithms being able to process billions of training data points, characterized by millions of features, trained over many days in order to compute accurate weights for each node in each layer.

4.2.3 Recurrent Neural Networks for Deep Learning

Unlike traditional ANNs that are feedforward networks, recurrent neural networks (RNNs) have a recurrent feedback connection [32], [137]. These networks learn patterns in the input dataset by modifying weights calculated from the previous state stored and can hence process arbitrary sequences of inputs by allowing persistence in information. They hold two major advantages over ANNs. First, they can process arbitrary lengths of sequence inputs. This is useful for pattern recognition in time series data as the same amount of information is not always available at each epoch. Second and more importantly, they can share information learned from features across different points in time and are relatively invariant to the exact point of occurrence of these features. Hence, they are designed to focus more on temporal dependencies of the features of a datapoint when forming associations with the class label. These networks are most often used to learn patterns from word groupings in natural language processing (NLP) applications, where the outcome can be to translate from one language (input) to another language (output), or to analyze the
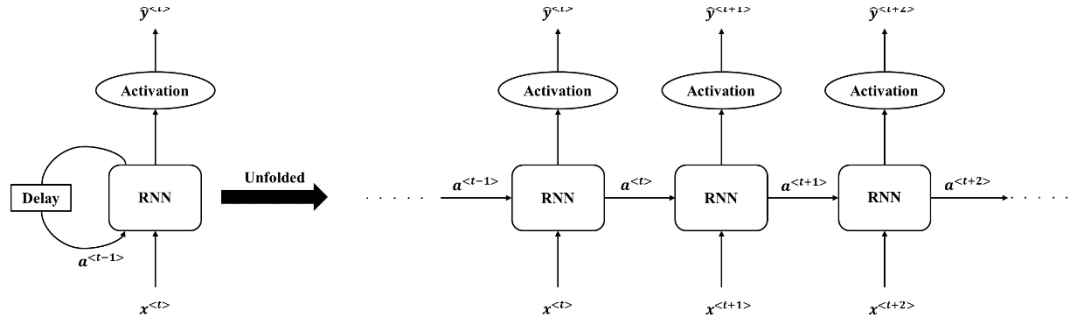
Figure 4.2: Unrolled representation of a recurrent neural network (RNN)

sentiment of a given set of words such as assigning a rating out of five stars automatically for a movie based on the sentiment of the review. However, these networks can be used to learn patterns from any dependent sets of inputs.

Figure 4.2 shows an "unrolled" representation of a recurrent neural network, where the computations across time can be visualized as different representations of the same unit. The information transmitted from one RNN to the next is represented by the activation at time t as $a^{<t>}$, the output of the RNN at time t written as $\hat{y}^{<t>}$, with the corresponding input as $x^{<t>}$. A one-hot sparse encoding of $x^{<t>}$ is used, calculated from a dictionary stored in the database. The equations to determine $a^{<t>}$ and $\hat{y}^{<t>}$ are as follows:

$$a^{<t>} = g_1(\boldsymbol{w}_{aa}a^{<t-1>} + \boldsymbol{w}_{ax}x^{<t>} + b_a) \qquad (4.8)$$

$$\hat{y}^{<t>} = g_2(\boldsymbol{w}_{ya}a^{<t-1>} + b_y) \qquad (4.9)$$

Unique weight vectors $\mathbf{w}_{aa}$ and $\mathbf{w}_{ax}$ are computed for the activations from the previous layer and the current input, respectively. The weight matrix $\mathbf{w}_{aa}$ can be thought of as the importance assigned to the activations calculated from the previous layer, while $\mathbf{w}_{ya}$ resembles the weight vector computed for the traditional ANN. Activation functions g_1 and g_2 are non-linear functions resembling their counterparts in the ANN. Activation function g_1 is usually a tanh function. Since there are many RNNs chained together in time, backpropagation can cause the gradient to vanish. The tanh function can combat this effectively due to the larger gradient near zero, while the sigmoid has a gradient close to zero for small values. However, the sigmoid function is widely used as the activation function for g_2. In order to simplify notation, (4.8) can be reformulated as:

$$a^{<t>} = g_1(\mathbf{w}_a[a^{<t-1>}, x^{<t>}] + b_a) \qquad (4.10)$$

RNNs can either calculate the output at every iteration, every few iterations, or just one output at the end of all iterations. The latter configuration is useful to the problem of pain detection, as the temporal dependencies between microstates relevant to pain can be learned from their stability and frequency in the corresponding microstate sequence.

The loss function computed for each RNN unit is summed across all units, and is depicted as follows:
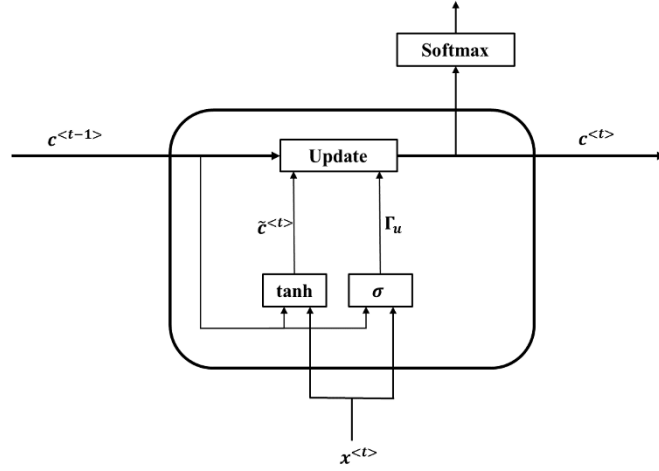
Figure 4.3: Visualization of a gated recurrent unit (GRU)

$$L^{<t>}(\hat{y}^{<t>}, y^{<t>}) = -y^{<t>}ln(\hat{y}^{<t>}) - (1 - y^{<t>})ln(1 - \hat{y}^{<t>}) \qquad (4.11)$$

$$L(\hat{y}, y) = \sum_{t=1}^{T} L^{<t>}(\hat{y}^{<t>}, y^{<t>}) \qquad (4.12)$$

In theory, while RNNs are capable of learning long-term dependencies, learning the relationship between two features α and β becomes less effective as the time gap between them grows larger due to successive backpropagation across several RNN units in time. While this can be somewhat mitigated by carefully choosing initial conditions over the randomized distributions of **w** for gradient descent, these initial values would be problem-specific, and do not generalize to similar, independently collected datasets.

4.2.4 Learning Long-term Dependencies with Gated Recurrent Units

Gated Recurrent Units (GRU) were developed in [137] to minimize the vanishing gradient problem. A typical GRU unit is shown in Figure 4.3. The major difference between a GRU and a standard RNN is the addition of an "update" gate that is a measure of importance assigned to activations from all previous layers. The higher the value of the update factor Γ_u, the more information will be retained from previous layers.

More formally, at time t, the activation $c^{<t-1>}$ is modified as follows:

$$\tilde{c}^{<t>} = tanh(\mathbf{w}_c[c^{<t-1>}, x^{<t>}] + b_c) \tag{4.13}$$

$$\Gamma_u = \sigma(\mathbf{w}_u[c^{<t-1>}, x^{<t>}] + b_u) \tag{4.14}$$

$$c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + (1 - \Gamma_u) * c^{<t-1>} \tag{4.15}$$

$$\hat{y}^{<t>} = softmax(c^{<t>}) \tag{4.16}$$

The tanh function is used to combat the vanishing gradient problem. The variable $\tilde{c}^{<t>}$ is learned based on the activations from the previous layers $c^{<t-1>}$ and the current input $x^{<t>}$. The update gate Γ_u is learned in a similar fashion, but the weight matrix $\mathbf{w}_u$ is understood to learn parameters specifically relating to the importance of $c^{<t-1>}$. The update gate is then used to weigh both $\tilde{c}^{<t>}$ and $c^{<t-1>}$ before the final activation $c^{<t>}$ is passed to the next layer.

A more general variant of the GRU, called the Long Short-term Memory (LSTM) network has been traditionally used for learning from sequences. LSTMs differ from the GRU by splitting Γ_u into Γ_r (recall) and Γ_f (forget) gates, each with their unique weight matrices. Modern RNNs prefer the use of GRUs compared to LSTMs due to significant savings in computational time, with only a small drop in performance.

4.2.5   Overview of Approach

The aim of this chapter is to demonstrate a deep learning algorithm that can effectively detect pain using a pain-specific spatiotemporal biomarker. This chapter also details efforts towards detecting the onset of pain in the presence of a non-painful, task-specific confound in order to demonstrate the generalizability and specificity of the pain response and the associated deep learning model.

4.3   Methods

4.3.1   Data collection

The dataset collected in Chapter 3 was used for objective pain detection. In short, 27 subjects were recruited to collect EEG data related to pain, non-painful heat, task-related attention, and resting state. The thermal stimulus was applied to the underside of the right forearm. The continuous N-back task was used to collect EEG data

related to attention. Each subject underwent a training session on the morning of the study, where they learned to recognize the difference between pain and non-painful heat (pain threshold) and learned to identify the most pain they were willing to bear (pain tolerance). EEG data were recorded on the afternoon of the study corresponding to pain, resting state, non-painful heat and task. In addition, EEG data corresponding to each subject identifying their pain threshold and pain tolerance were also collected in the afternoon session.

An additional paradigm was incorporated into the afternoon session, where the subjects performed 6 trials of the continuous N-back task for one minute. For each trial, starting at a random time point, a thermal stimulus lasting 10 seconds was administered. The subject was instructed to continue to perform the task to the best of their ability throughout the one-minute period. The temperature of the thermal stimulus was known to induce either non-painful heat, or pain almost at the tolerance of the subject. The goal of this paradigm was to obtain EEG data corresponding to two stimuli that consistently activated the same brain regions that produced two different temporal responses.

4.3.2 Preprocessing and Microstate Analysis

For all trials, EEG data were recorded at 1024 Hz using a 128-channel Biosemi Active-two system. The offset was kept below 20 mV for all conditions. The

continuous rating was collected at a rate of 192 Hz. During the analysis of each set of pain/heat trials, the first 1.5 seconds of EEG data were removed to minimize the effects of salience due to the drastic change in stimulus intensity. The first trial for each set of task trials was also removed to eliminate effects of large hand/eye movements. Subsequent analysis was performed in MATLAB and R using native toolboxes and those developed by EEGLAB [60].

The EEG signals were filtered and down sampled to 256 Hz and band-pass filtered between 1 Hz and 100 Hz to remove any muscle contraction artifacts and voltage drifts. Line noise at 60 Hz was filtered out using a notch filter, and further filtered and down sampled to 100 Hz. EEG data were then re-referenced to average reference. Artifact subspace reconstruction [61] was used to clean eye blinks and reject bad time windows. Ocular and muscle artifacts in EEG were removed using blind source separation [115].

For each subject, the microstates were computed according to the temporal approach detailed in [131]. For this analysis, a dictionary of 30 microstates and their corresponding fits were calculated for the instantaneous EEG scalp topography map per time point. Each spatial map was then assigned a fuzzy label with information collected from all 30 clusters.

4.3.3   Embedding of EEG Scalp Maps

As mentioned in section 4.2.3, the most common application of RNNs has been in natural language processing (NLP), which the field of analyzing patterns in spoken/written language. These networks evaluate the relationships between groups of words in a sequence for applications such as word/phrase prediction, sentiment classification, and trigger-word detection. For such applications, a computationally inefficient approach would be to encode each word as a one-hot sparse vector acting as a pointer to a position in a precomputed dictionary. This orthogonal representation is not meaningful as it provides no information on the dependencies between the words. Rather, relationships between words can be encoded using word embeddings, which are dense vector representations of the inputs. The main rationale behind creating embeddings is to use continuous vectors for training, on which backpropagation in RNNs works more efficiently.

Some common embeddings can be created using nearest neighbor algorithms and cosine similarities. Traditional word vector representations [138] have been created for NLP applications using a separate neural network trained on a large database of words in order to learn embeddings for a particular application. However, such an approach for encoding EEG spatial maps would require millions of training examples, which was not available for the current dataset. Another way of creating

Output

Batch normalization

Dropout layer

GRU layer

x_1　　　　x_2　　　　x_3　　　　　　　x_n

Dictionary of microstate fits

$$\begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ . \\ . \\ . \\ x_{1p} \end{bmatrix} \quad \begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \\ . \\ . \\ . \\ x_{2p} \end{bmatrix} \quad \begin{bmatrix} x_{31} \\ x_{32} \\ x_{33} \\ . \\ . \\ . \\ x_{3p} \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \begin{bmatrix} x_{n1} \\ x_{n2} \\ x_{n3} \\ . \\ . \\ . \\ x_{np} \end{bmatrix}$$
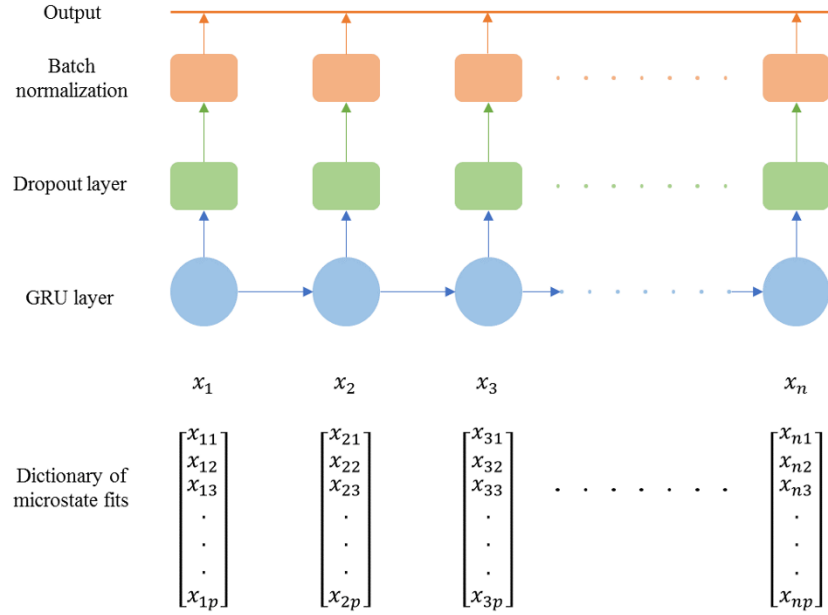
Figure 4.4: Components of the main deep learning (DL) layer

such embeddings would be to use microstates calculated using unsupervised clustering techniques to create a dictionary of fuzzy microstate labels corresponding to each datapoint rather than using the dominant microstate, as explained in section 4.1. This embedding can then be used to better model the temporal dependencies between spatially distinct regions of activation to be used as inputs for the GRU units.

4.3.4　GRU Network Architectures for Pairwise Comparisons

There were three main deep learning layers stacked together to form the core deep learning (DL) layer, as shown in Figure 4.4. The GRU layer, explained in section 4.2.4 was used as the main functional cell. The number of hidden states, or

the activation units ($c^{<t>}$) is a configurable hyperparameter and determined during hyperparameter tuning. The batch-normalization layer standardizes the outputs of the GRU and the ANN layers to minimize bias due to covariance shift after the computations in each hidden layer, before activation. The dropout layer is an effective regularization layer, where information from a randomized subset of hidden nodes is discarded to prevent overfitting, which is an unfortunately common occurrence during training on EEG data, due to high correlation between time points. The percentage of hidden nodes to discard at each training epoch is also trainable. The activation functions for all hidden units in the GRU was the tanh function to minimize the vanishing gradient problem.

Due to the various frequency bands implicated in each of the four conditions (pain, non-painful heat, task and rest), the pain detection algorithm was implemented using a divide-and-conquer approach. Specifically, three network architectures were constructed and trained on pairwise conditions (pain vs heat, pain vs task and pain vs rest). For pain vs heat, the EEG data were filtered in the high beta-low gamma frequency bands (20-50 Hz), as these were most active for somatosensory stimuli [28], [52]. For pain vs task, the EEG data were filtered between the theta-low beta bands (4-20 Hz) according to past research that has shown these bands to be widely active during attention to different task-related stimuli [111], [112]. The activity in all available frequency bands were used for the pain vs rest paradigm.
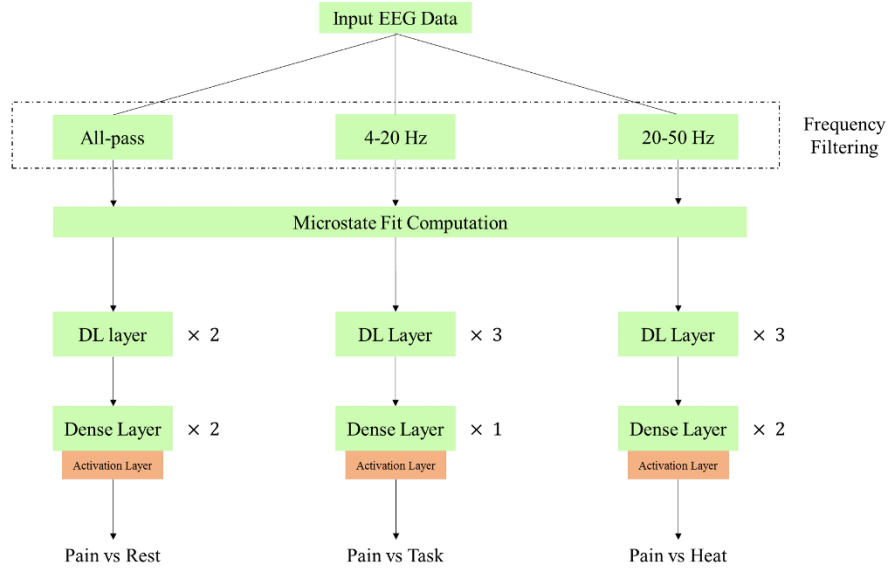
Figure 4.5: Flow diagram of the training paradigm

4.3.5   Training, Validation, and Testing Paradigms

The training paradigm is shown in Figure 4.5. All deep learning architectures were trained using a NVIDIA GeForce GTX 745 GPU on an Intel Core i7-4790 CPU at 3 GHz. The input EEG data was filtered in the frequency spectrum as detailed. The microstate clusters and fits were computed according to [131], which were then epoched into 600 samples and propagated through the custom-tuned deep learning architectures designed for each condition pair. The dense layers are standard feed-forward ANN layers, where the number of hidden nodes is trainable.
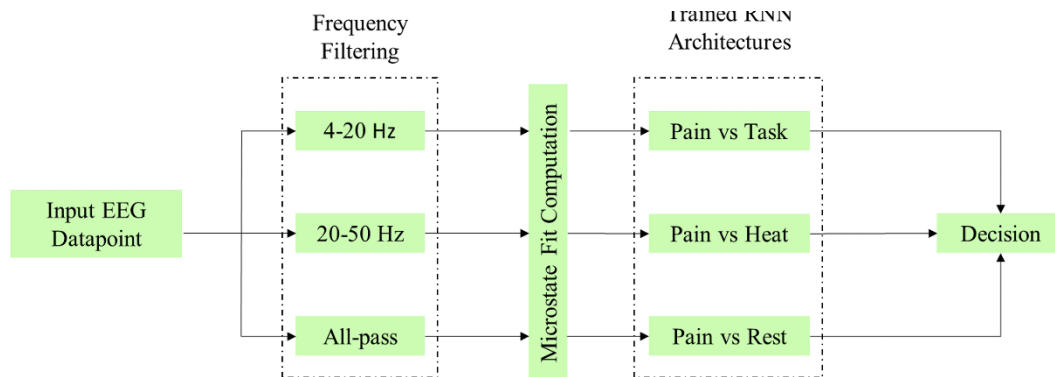
Figure 4.6: Flow diagram of testing paradigm

For training, 80% of the trials from 22 subjects corresponding to non-painful heat, pain, resting state and task-related attention were used, while cross validation was carried out on the remaining 20% of the trials, in addition to two of the most consistent tolerance trials, and the trials from the second paradigm of the experiment (N-back task interrupted by thermal stimulus). Hyperparameter optimization during cross-validation for each of these architectures was performed using tree-structured parzen estimators (TPE) to determine the number of hidden nodes, the parameters of each layer etc. In short, this optimization algorithm works by creating multiple randomized subsets of the hyperparameter space to train the network using training data, which then provides a score on the validation data. The subsets corresponding to the top 10-25% of the cross-validation scores are then assumed to contain the optimal hyperparameters, creating two groups in the union of the hyper-parameter

subsets. The next step in the TPE algorithm is to model the likelihood probability for each of the two groups using parzen-window density estimators. Obtaining $l(x)$ for the first group and $g(x)$ for the second group, the expected improvement (EI) is given by:

$$EI(x) = \frac{l(x)}{g(x)} \tag{4.17}$$

The top 10-25% of the hyperparameter subsets with the highest EI are then used in the next iteration, until there is only one set remaining, reaching convergence.

The trained network was then tested on five unseen test subjects, as shown in Figure 4.6. The steps to prepare the EEG data for feeding into the deep learning architectures were similar to those during the training/validation phase. The filtered, epoched segment of EEG data is then run through each of the three network architectures. Based on the confidence scores generated by each binary network, the final decision of "pain" or "no pain" is made.

4.3.6 Comparisons to Previous Algorithms

The performance of this pain detection algorithm was compared to the pain quantification algorithm developed in [24] in order to verify robustness and specificity to pain. The pain quantification algorithm is used here as a binary classifier, where the random forest (RF) model is trained to predict "pain" vs "no

pain". All EEG responses to resting state, task-related attention, and non-painful thermal stimuli were labelled as "no pain" for the pain quantification algorithm. The training and testing datasets are similar to the one detailed in section 4.3.5. The comparisons were made for resting state vs pain, task-related attention vs pain and non-painful heat vs pain. An additional comparison was also conducted for both non-painful attention paradigms vs finer resolutions of the continuous pain score collected during the painful trials. The continuous score was uniformly binned into discrete pain scores of 1-3.

4.3.7   Combining the Pain quantifier and the Pain detector

Finally, the RF-based pain quantification algorithm is stacked below the deep learning pain detection algorithm to robustly detect and quantify pain, in a two-stage process. The first stage is the deep learning algorithm that can robustly detect pain, passing the predicted "pain" segments to the second stage, which is the RF-based pain quantification algorithm that can effectively quantify pain given the presence of pain. The results of this two-stage pain quantifier are then compared to the performance of the RF-based pain quantifier implementation alone, to investigate any improvements made to the previous state of the art.

4.4 Results

The plot for the training and validation loss for the pain vs heat branch is shown in Figure 4.8. The training accuracy before convergence was 91.38%, while the validation accuracy was 88.93%. A similar plot for the pain vs task branch, shown in Figure 4.9, resulted in a training accuracy 94.19% and a validation accuracy of 90.66%. The best results were in the pain vs rest branch as seen in Figure 4.7, where the training accuracy was 98.38% and the validation accuracy was 96.72%. The steady drop in both the training the validation losses indicates that sufficient regularization has been implemented for each architecture, with the algorithm implementing early stopping before overfitting can occur. The complete list of tuned hyperparameters for each layer in each architecture is shown in Table 4.1.

A comparison of performances between the deep learning pain detector and the RF-based pain quantifier is shown in Figure 4.10 for both rest vs pain and non-painful attention vs pain. The plots compare the accuracy and the specificity with respect to correct classification of non-painful attention. The "non-painful" label consists of EEG data corresponding to both task-related attention and non-painful thermal stimulation. The accuracies are comparable, with the rest vs pain using the deep learning pain detector scoring 96.49 ± 1.13 %, while it is 96.07 ± 1.65 % for the RF-based pain detector. Specificities are also similar to each other. However, the difference in performance is clear during the non-painful attention vs pain paradigm.

Both in terms of accuracy ($92.55 \pm 1.12\%$) and more importantly specificity (90.67 ± 1.42 %), the deep learning algorithm outperforms the RF-based pain detector (83.10 ± 3.06 % and $76.47 \pm 2.88\%$, respectively). This implies that more non-painful data segments are correctly classified, satisfying the objective of this work, which was to construct a pain-specific machine learning detector that can be deployed outside of experimental conditions and clinical settings.

A closer look at the classification results for the non-painful attention paradigm is shown in Figure 4.11. In this analysis, task-related attention and non-painful heat are compared separately with pain to compare changes in specificity. The lowest performance is seen consistently in the RF-based classifier for task-related attention vs pain. Task vs low pain has the biggest contrast in performance with $77.60 \pm 5.38\%$, while it was 86.97 ± 4.73 % using the pain detector developed in this chapter. Hence, task-related attention is the biggest detriment to the RF-based pain detector, with the biggest confound being EEG data associated with low-pain scores. This is remedied by using the trained deep learning algorithm, which outperforms the latter algorithm in all discrete pain classes. Improvements in performance is also seen in the non-painful heat vs pain paradigm using the deep learning algorithm, as shown in Figure 4.11.

The results obtained above are emphasized in Figure 4.12 by comparing the per-class F-score of the two-stage pain quantification algorithm with the single-stage

random forest-based pain quantifier developed previously, for the pain scores described in section 4.3.6. The per-class F-score, which is the harmonic mean of the precision and recall, is a complementary metric to specificity for assessing performance in pain detection and quantification. Using the two-stage quantifier noticeably improves the F-score, especially in the lower end of the pain scale. The major contribution to increase in performance is due to the first pain detection stage, as demonstrated by an F-score of $91.628 \pm 3.016\%$ for the two-stage quantifier compared to $81.657 \pm 4.996\%$ for the RF-based quantifier, a 10% improvement. The subsequent differences in F-score between the other classes can be attributed to more frequent misclassifications of the "no pain" class for the RF-based quantifier.

4.5 Discussion and Conclusions

This chapter outlines a robust pain detector that improves on the previously developed pain quantification algorithm. This detector uses gated recurrent neural nets to construct a deep learning architecture for classifying "pain" vs "no pain", with attention given to increasing the specificity of the algorithm. The detector minimizes the limitations of state of the art pain quantification algorithms and can be coupled with it to make pain quantification more robust. In addition to added specificity, the current two-stage pain detection/quantification tool has the best and most robust performance among existing EEG quantification algorithms. The ultimate objective

of this tool is to further the potential of a pain quantification tool to be used outside of a clinic by minimizing the number of false alarms related to pain.

The algorithm is designed as a three-pronged deep learning paradigm. Each branch of the main deep learning algorithm is a custom-tuned binary neural network that is trained specifically to distinguish between pain and a non-painful experimental condition. Due to the high temporal resolution of EEG, frequency bands known to be active for certain non-painful conditions are preserved before propagating the data through the deep net, while the others are filtered out. This architecture relies on unique patterns in the distributions of temporal dependencies between instantaneous EEG spatial maps which have been demonstrated to be specific to pain using a different formulation outlined previously.

The most significant advantage is in its generalizability, as more non-painful conditions outside those in this work can simply be trained concurrently by adding another neural network branch to the main pain detection algorithm. This branch can be trained separately to distinguish between pain and the non-painful condition under investigation. In addition, this architecture can be stacked on top of a pain quantification algorithm to improve its performance. The initial results illustrated in this chapter demonstrates the generalizability of the extracted temporal pain response and the pain detector. Such a tool can hence be used to track, record and compare the

onset and duration of a painful episode to previously recorded episodes for evaluating the degree of success of pain therapies over time.

One limitation of this approach is the lack of interpretability of the derived pain response. The use of a dictionary of microstates to exploit the temporal evolution of pain prevents an investigation into which microstate most informs the pain detection algorithm. In addition, while neural networks are capable of automatic feature abstraction, it is difficult to attribute neuroscientific significance to how exactly the microstate activations are weighted. With current deep learning research progressing towards interpretability of layer-to-layer weights [139], further investigations in pain research following this school of thought would likely yield more information on how the brain processes pain.

Table 4.1: List of parameters for each deep learning layer and architecture

Layer name	Hyper-parameter	Pain vs Rest	Pain vs Heat	Pain vs Task
Deep learning layer 1	Number of units	100	10	100
	Dropout probability	0.136458	0.5642	0.7645
Deep learning layer 2	Number of units	50	100	50
	Dropout probability	0.64265	0.31657	0.67416
Deep learning layer 3	Number of units	-	100	50
	Dropout probability	-	0.69732	0.4628
Dense layer 1	Number of nodes	20	-	20
	Activation function	tanh	-	tanh
Dense layer 1	Number of nodes	2	2	2
	Activation function	softmax	softmax	softmax
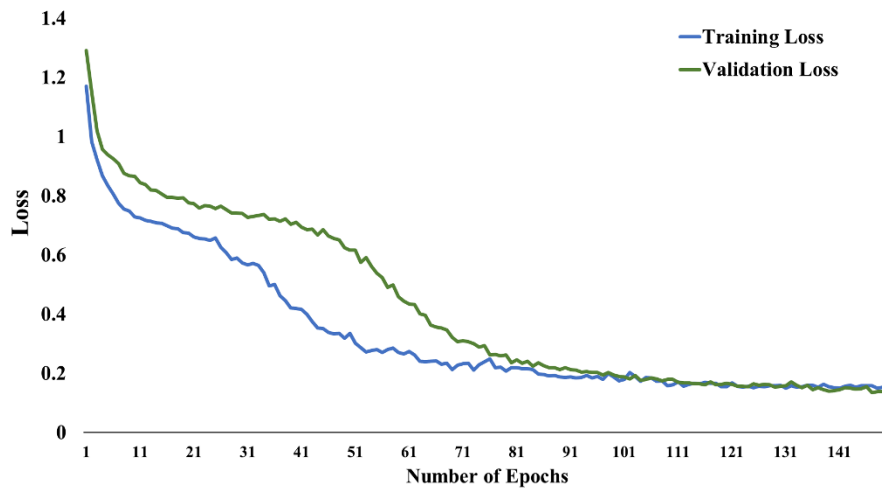
Figure 4.8: Plot of training and validation losses for the pain vs rest classifier branch
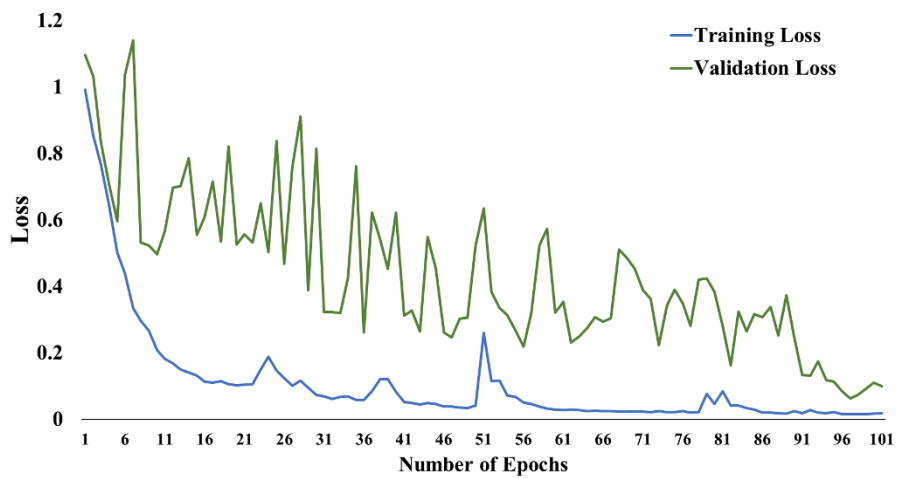
Figure 4.7: Plot of training and validation losses for the pain vs heat classifier branch
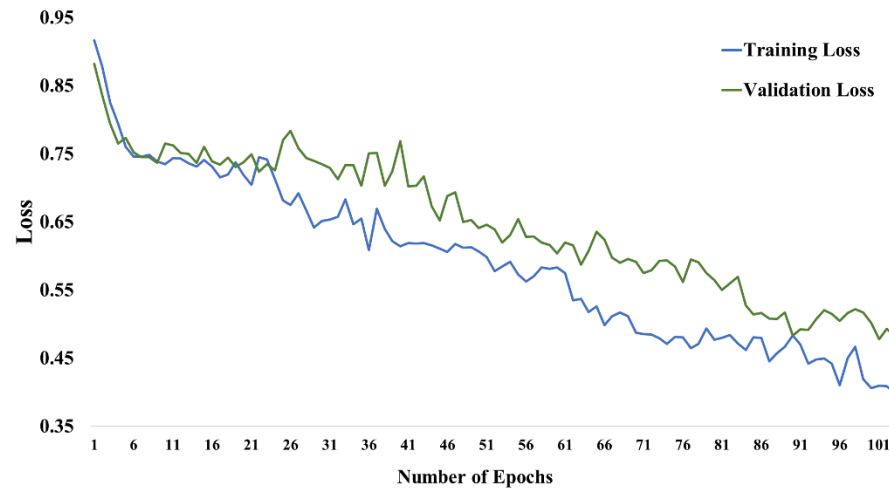
Figure 4.9: Plot of training and validation losses for the pain vs task classifier branch
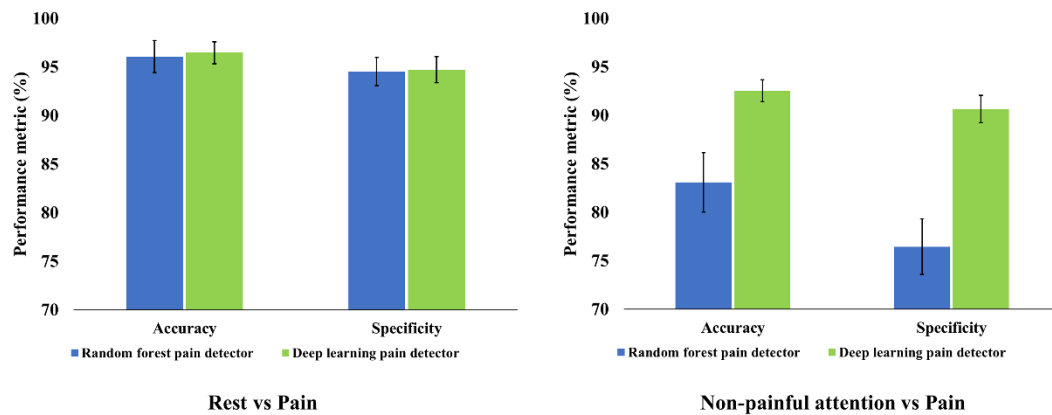
Figure 4.10: Performance metrics for comparing the deep learning pain detector and the random forest pain quantifier
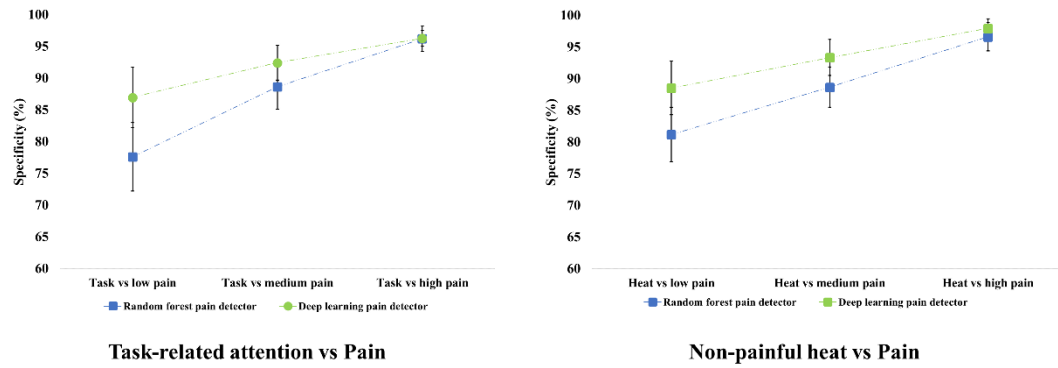
Figure 4.11: Comparison of specificities of the two pain detectors to detect different levels of pain from task-related attention and non-painful heat
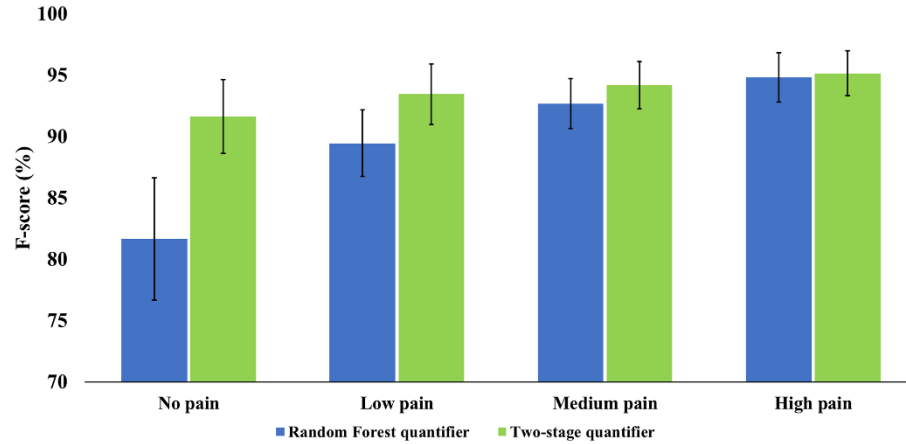
Figure 4.12: Comparison of per-class F-score for the single stage RF pain quantifier, and the two-stage pain quantifier

Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis has focused on the development of robust brain biomarkers that are specific to thermal pain using noninvasive electroencephalography (EEG), and the design of new machine learning and deep learning pipelines to objectively detect and quantify pain using these biomarkers. Investigations into the temporal interactions between spatially distinct brain regions have yielded patterns of activity that are more specific to pain than previously observed. The machine learning algorithms use these temporal dependencies to improve on the current state of the art in terms of accuracy, robustness, and specificity to pain. The research carried out in this work aims to translate the last five decades of pain research into fast and accurate detection and characterization of pain in the clinic to improve the quality of life of the individual. After a reliable and objective baseline on pain has been established, the clinician can make a treatment recommendation based on their pain history, current pain state, as well as taking into account their subjective pain scores for reference.

5.1.1 Quantification of Thermal Pain

The first algorithm presented was used to quantify pain during the administration of a noxious thermal stimulus. A new study was designed to collect EEG data on 25 human subjects where a tonic thermal stimulus was administered to collect the continuous pain score ranging from the subject's pain threshold (lower bound on pain) and their pain tolerance (upper bound on pain). The focus was on obtaining consistent ground truth pain scores with the range of stimulus intensity curated for each subject. The temporal wavelet spectrum of independent components was used to characterize the EEG information corresponding to the pain score and train a random forest model. The most informative EEG spatial map and the best performing model were found during cross validation and used for testing on the remaining subject. The temporal activity of the best performing component was also evaluated in the spatial and the frequency domains. The pain quantification algorithm achieved mean accuracy of 95.33% for 2-way classification (rest vs pain) and 89.45% accuracy in 10-way pain quantification, where the pain score ranged from 1 (no pain) to 10 (maximum tolerable pain). The present approach outperforms existing state of the art methods for pain quantification by reliably predicting the pain score and demonstrating scalability of the approach for finer resolutions of a continuous pain scale, suggesting the generalizability of a biomarker for pain across a population. The gamma band in EEG is found to be the most important predictor of pain for all

resolutions of the pain score across inter-subject and intra-subject prediction paradigms. In addition, it was shown that the differences between the shared neural correlates of the pain score across populations and the unique signature at an individual level are observable, hence can be quantified in further explorations by extending these approaches; unraveling a potential pain response mechanism.

5.1.2 Detection of Pain-specific Neurological Responses

In order to verify the efficacy of a given pain therapy procedure in clinical trials for pain patients, an automated algorithm that can detect and record the intensity painful episodes outside the clinic would be very informative to clinicians in evaluating the next steps in clinical research. With increasing research efforts spent on increasing the affordability and convenience of interfacing neuroimaging technology using EEG with other devices, an objective method of detecting and tracking pain can become a viable option in the near future. However, the current state of the art focuses on detecting pain by contrasting most commonly against resting state, where the subject tries to remain motionless and let their mind wander for the duration of the EEG recording. Given that the brain regions participating in pain processing are non-specific to pain, such algorithms would produce many false alarms in the real world where the subject is engaged in performing non-painful tasks, and record brain responses not specific to painful episodes. This thesis tackled this issue at the feature engineering level by analyzing the interaction between brain

regions by investigating the temporal pattern of quasi-stable, topographical maps in EEG called microstates. A pattern was observed that was unique to pain, which could differentiate between other stimuli known to evoke similar spatial responses. The specificity of this pattern was statistically verified by analyzing the differences in the temporal microstate sequences extracted from experimental stimuli designed to capture attentional, somatosensory and painful correlates of stimulus processing. It was observed that the trends of first order microstate parameters such as microstate duration and occurrence, while significantly different for non-rest conditions compared to rest, were too chaotic to be specific to pain. Quantifying the heavy-tailed nature of the autocorrelation of the temporal microstate sequence by modelling the interactions of these sequences across increasing time scales using wavelet leaders illustrated the presence of microstate behavior unique to pain. These results signified a step forward in effective characterization of a unique temporal pain response that can be used for pain detection and quantification.

5.1.3 Deep Learning for Objective Pain Detection

Building on all of the results detailed above, a fast pain detection algorithm was developed using deep learning techniques. A generalization of the temporal microstate sequence was used, where the instantaneous EEG spatial map was characterized by a dictionary of microstate fits, rather than using the dominant microstate to explain EEG activity at that time point. This facilitated robust pain

detection by analyzing non-orthogonal similarities in microstate fit between EEG data points across time scales using gated recurrent units (GRU), which are efficient implementations of neural networks designed to identify patterns in causal time series data. The deep learning network was developed as a tree classifier, with a binary GRU-based deep learning architecture forming each branch. Each branch was trained to differentiate between pain and a non-painful condition in order to increase the specificity of the final algorithm to pain. This has the advantage of training multiple branches in a parallel fashion as more non-painful responses to task-related stimuli are considered. Domain knowledge of said non-painful condition can also be incorporated by filtering in the specific frequency bands known to show the most activity when contrasting with pain. The resulting algorithm improved on the state of the art by scoring 90.67% in terms of specificity to pain, compared to both attentions to non-painful heat stimuli and particularly demanding memory-based N-back tasks. Stacking this deep learning pain detection algorithm on top of the pain quantification algorithm developed earlier improved on robustness, specificity and the accuracy of state of the art pain quantification algorithms. This is an encouraging step towards the research and development of tools for characterizing and quantifying pain, which can be used both within and outside the clinic to detect, characterize and track the progression of pain in patients.

5.2 Future Directions

Pain is defined as an inherently subjective sensation. Due to the temporal integration of spatially distinct responses in the brain, it is possible that the spatiotemporal correlates of affective and emotional processing that result from a nociceptive stimulus are uncommon across individuals. In the pain quantification algorithm, the temporal evolution of the most informative independent component is used. While it may seem consistently accurate, it is possible that the uniqueness of the pain response topology diminishes as more subjects are considered for analysis, and this may account for a drop in the accuracy when more subjects are considered for training. An extension of this pain quantification algorithm would be to analyze specific intervals of the responses to the tonic thermal stimulus, in order to determine when classification accuracy starts to deteriorate relative to the entire trial. This may shed more light into subject-specific pain processing, and lead to improvements in the current pain quantification algorithm

While developing the pain-specific temporal response using microstates, the results found by analyzing microstate behavior across time scales warrant further research. Specifically, investigations into the correlation decay across time scales, and conducting source-imaging studies on the EEG data at the time scale where maximum decorrelation was observed would shed light on the behavior of individual microstate dependencies. Thus, stronger ties can be established between temporal

patterns at the sub-second scale to the behavior of resultant resting state networks, increasing the neuroscientific significance of microstates in characterizing pain-specific responses.

The deep learning algorithm developed in this work can be improved on by training more branches to classify other types of non-painful responses to attention-related stimuli. In addition, it has been shown that engagement in other tasks can also modulate the experience of pain [12]. Research into the interaction and disruption of the task-specific neurological signature by a painful distractor stimulus can further improve the robustness of the pain detection algorithm. In addition, further research can be conducted into the interpretability of the temporal pain signature by analyzing the distribution of layer-to-layer weights in each branch of the deep learning architecture to get an idea of the differences in stimulus-specific processing patterns in the human brain.

References

[1]   I. Tracey and P. W. Mantyh, "The Cerebral Signature for Pain Perception and Its Modulation," *Neuron*, vol. 55, no. 3, pp. 377–391, Aug. 2007.

[2]   T. H. Lindsay *et al.*, "Pancreatic cancer pain and its correlation with changes in tumor vasculature, macrophage infiltration, neuronal innervation, body weight and disease progression," *Pain*, vol. 119, no. 1, pp. 233–246, Dec. 2005.

[3]   G. Blackburn-Munro, "Pain-like behaviours in animals – how human are they?," *Trends Pharmacol. Sci.*, vol. 25, no. 6, pp. 299–305, Jun. 2004.

[4]   C. M. Jones, K. A. Mack, and L. J. Paulozzi, "Pharmaceutical Overdose Deaths, United States, 2010," *JAMA*, vol. 309, no. 7, pp. 657–659, Feb. 2013.

[5]   Paulozzi Leonard J., Budnitz Daniel S., and Xi Yongli, "Increasing deaths from opioid analgesics in the United States," *Pharmacoepidemiol. Drug Saf.*, vol. 15, no. 9, pp. 618–627, Aug. 2006.

[6]   G. A. Cecchi *et al.*, "Predictive Dynamics of Human Pain Perception," *PLOS Comput. Biol.*, vol. 8, no. 10, p. e1002719, Oct. 2012.

[7]   D. J. Gaskin and P. Richard, "The Economic Costs of Pain in the United States," *J. Pain*, vol. 13, no. 8, pp. 715–724, Aug. 2012.

[8]   A. V. Apkarian, M. C. Bushnell, R.-D. Treede, and J.-K. Zubieta, "Human brain mechanisms of pain perception and regulation in health and disease," *Eur. J. Pain*, vol. 9, no. 4, pp. 463–463, Aug. 2005.

[9]   S. W. G. Derbyshire, "Exploring the pain 'neuromatrix,'" *Curr. Rev. Pain*, vol. 4, no. 6, pp. 467–477, Dec. 2000.

[10]   G. D. Iannetti and A. Mouraux, "From the neuromatrix to the pain matrix (and back)," *Exp. Brain Res.*, vol. 205, no. 1, pp. 1–12, Aug. 2010.

[11]   L. Hu and G. D. Iannetti, "Painful Issues in Pain Prediction," *Trends Neurosci.*, vol. 39, no. 4, pp. 212–220, Apr. 2016.

[12]   D. M. Torta, V. Legrain, A. Mouraux, and E. Valentini, "Attention to pain! A neurocognitive perspective on attentional modulation of pain in neuroimaging studies," *Cortex*, vol. 89, pp. 120–134, Apr. 2017.

[13]   H. Berger, "Über das Elektrenkephalogramm des Menschen," *Arch. Für Psychiatr. Nervenkrankh.*, vol. 87, no. 1, pp. 527–570, Dec. 1929.

[14]   B. He, L. Yang, C. Wilke, and H. Yuan, "Electrophysiological Imaging of Brain Activity and Connectivity #x2014;Challenges and Opportunities," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 7, pp. 1918–1931, Jul. 2011.

[15] R.-D. Treede, J. Lorenz, and U. Baumgärtner, "Clinical usefulness of laser-evoked potentials," *Neurophysiol. Clin. Neurophysiol.*, vol. 33, no. 6, pp. 303–314, Dec. 2003.

[16] B. J. Edelman, B. Baxter, and B. He, "EEG Source Imaging Enhances the Decoding of Complex Right-Hand Motor Imagery Tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4–14, Jan. 2016.

[17] V. Menon and L. Q. Uddin, "Saliency, switching, attention and control: a network model of insula function," *Brain Struct. Funct.*, vol. 214, no. 5–6, pp. 655–667, Jun. 2010.

[18] K. H. Brodersen *et al.*, "Decoding the perception of pain from fMRI using multivariate pattern analysis," *NeuroImage*, vol. 63, no. 3, pp. 1162–1170, Nov. 2012.

[19] E. G. Duerden and M.-C. Albanese, "Localization of pain-related brain activation: A meta-analysis of neuroimaging data," *Hum. Brain Mapp.*, vol. 34, no. 1, pp. 109–149, Jan. 2013.

[20] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 14–30, Nov. 2001.

[21] H. Vogel, J. D. Port, F. A. Lenz, M. Solaiyappan, G. Krauss, and R.-D. Treede, "Dipole Source Analysis of Laser-Evoked Subdural Potentials Recorded From Parasylvian Cortex in Humans," *J. Neurophysiol.*, vol. 89, no. 6, pp. 3051–3060, Jun. 2003.

[22] B. He, "Brain electric source imaging: scalp Laplacian mapping and cortical imaging.," *Crit. Rev. Biomed. Eng.*, vol. 27, no. 3–5, pp. 149–188, 1999.

[23] E. Schulz, L. Tiemann, T. Schuster, J. Gross, and M. Ploner, "Neurophysiological Coding of Traits and States in the Perception of Pain," *Cereb. Cortex*, p. bhr027, Mar. 2011.

[24] © 2018 IEEE. Reprinted, with permission, from V. Vijayakumar, M. Case, S. Shirinpour, and B. He, "Quantifying and Characterizing Tonic Thermal Pain Across Subjects From EEG Data Using Random Forest Models," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 12, pp. 2988–2996, Dec. 2017.

[25] G. Misra, W. Wang, D. B. Archer, A. Roy, and S. A. Coombes, "Automated classification of pain perception using high-density electroencephalography data," *J. Neurophysiol.*, vol. 117, no. 2, pp. 786–795, Feb. 2017.

[26] M. N. Baliki *et al.*, "Chronic Pain and the Emotional Brain: Specific Brain Activity Associated with Spontaneous Fluctuations of Intensity of Chronic Back Pain," *J. Neurosci.*, vol. 26, no. 47, pp. 12165–12173, Nov. 2006.

[27] T. D. Wager, L. Y. Atlas, M. A. Lindquist, M. Roy, C.-W. Woo, and E. Kross, "An fMRI-Based Neurologic Signature of Physical Pain," *N. Engl. J. Med.*, vol. 368, no. 15, pp. 1388–1397, Apr. 2013.

[28] E. Schulz *et al.*, "Prefrontal Gamma Oscillations Encode Tonic Pain in Humans," *Cereb. Cortex*, p. bhv043, Mar. 2015.

[29] C.-W. Woo, L. J. Chang, M. A. Lindquist, and T. D. Wager, "Building better biomarkers: brain models in translational neuroimaging," *Nat. Neurosci.*, vol. 20, no. 3, pp. 365–377, Mar. 2017.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[31] L. Zhao and C. E. Thorpe, "Stereo- and neural network-based pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 3, pp. 148–154, Sep. 2000.

[32] A. Graves, A. r Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6645–6649.

[33] A. Marquand, M. Howard, M. Brammer, C. Chu, S. Coen, and J. Mourão-Miranda, "Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes," *NeuroImage*, vol. 49, no. 3, pp. 2178–2189, Feb. 2010.

[34] A. J. Casson, S. Smith, J. S. Duncan, and E. Rodriguez-Villegas, "Wearable EEG: what is it, why is it needed and what does it entail?," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 5867–5870.

[35] D. Looney *et al.*, "The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring," *IEEE Pulse*, vol. 3, no. 6, pp. 32–42, Nov. 2012.

[36] J. M. Foss, A. V. Apkarian, and D. R. Chialvo, "Dynamics of Pain: Fractal Dimension of Temporal Variability of Spontaneous Pain Differentiates Between Pain States," *J. Neurophysiol.*, vol. 95, no. 2, pp. 730–736, Feb. 2006.

[37] M. Case *et al.*, "Characterization of functional brain activity and connectivity using EEG and fMRI in patients with sickle cell disease," *NeuroImage Clin.*, vol. 14, pp. 1–17, 2017.

[38] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Hum. Brain Mapp.*, vol. 2, no. 4, pp. 189–210, Jan. 1994.

[39] J.-D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nat. Rev. Neurosci.*, vol. 7, no. 7, pp. 523–534, Jul. 2006.

[40] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fMRI data," *Trends Cogn. Sci.*, vol. 10, no. 9, pp. 424–430, Sep. 2006.

[41] M. J. Rosa and B. Seymour, "Decoding the matrix: Benefits and limitations of applying machine learning algorithms to pain neuroimaging:," *Pain*, vol. 155, no. 5, pp. 864–867, May 2014.

[42] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning. Vol. 1. 2006*. Springer, New York.

[43] M. A. J. van Gerven, P. Kok, F. P. de Lange, and T. Heskes, "Dynamic decoding of ongoing perception," *NeuroImage*, vol. 57, no. 3, pp. 950–957, Aug. 2011.

[44] S. Klöppel, A. Abdulkadir, C. R. Jack Jr., N. Koutsouleris, J. Mourão-Miranda, and P. Vemuri, "Diagnostic neuroimaging across diseases," *NeuroImage*, vol. 61, no. 2, pp. 457–463, Jun. 2012.

[45] J. T. Giacino, J. J. Fins, S. Laureys, and N. D. Schiff, "Disorders of consciousness after acquired brain injury: the state of the science," *Nat. Rev. Neurol.*, vol. 10, no. 2, pp. 99–114, Feb. 2014.

[46] B. Bromm and A. C. N. Chen, "Brain electrical source analysis of laser evoked potentials in response to painful trigeminal nerve stimulation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 95, no. 1, pp. 14–26, Jul. 1995.

[47] M. Valeriani, D. Le Pera, D. Niddam, A. C. N. Chen, and L. Arendt-Nielsen, "Dipolar modelling of the scalp evoked potentials to painful contact heat stimulation of the human skin," *Neurosci. Lett.*, vol. 318, no. 1, pp. 44–48, Jan. 2002.

[48] L. Arendt-Nielsen and A. C. N. Chen, "Lasers and other thermal stimulators for activation of skin nociceptors in humans," *Neurophysiol. Clin. Clin. Neurophysiol.*, vol. 33, no. 6, pp. 259–268, Dec. 2003.

[49] C. Domnick, M. Hauck, J. Lorenz, A. Engel, and K. Casey, "C-fiber-related EEG-oscillations induced by laser radiant heat stimulation of capsaicin-treated skin," *Journal of Pain Research*, 17-Mar-2009. [Online]. Available: https://www.dovepress.com/c-fiber-related-eeg-oscillations-induced-by-laser-radiant-heat-stimula-peer-reviewed-article-JPR. [Accessed: 10-Jan-2017].

[50] R.-R. Nir, A. Sinai, E. Raz, E. Sprecher, and D. Yarnitsky, "Pain assessment by continuous EEG: Association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest," *Brain Res.*, vol. 1344, pp. 77–86, Jul. 2010.

[51] J. Gross, A. Schnitzler, L. Timmermann, and M. Ploner, "Gamma Oscillations in Human Primary Somatosensory Cortex Reflect Pain Perception," *PLOS Biol.*, vol. 5, no. 5, p. e133, Apr. 2007.

[52] L. Tiemann, E. Schulz, J. Gross, and M. Ploner, "Gamma oscillations as a neuronal correlate of the attentional effects of pain:," *Pain*, vol. 150, no. 2, pp. 302–308, Aug. 2010.

[53] E. Schulz, A. Zherdin, L. Tiemann, C. Plant, and M. Ploner, "Decoding an Individual's Sensitivity to Pain from the Multivariate Analysis of EEG Data," *Cereb. Cortex*, p. bhr186, Jul. 2011.

[54] M. T. Huber, J. Bartling, D. Pachur, S. v Woikowsky-Biedau, and S. Lautenbacher, "EEG responses to tonic heat pain," *Exp. Brain Res.*, vol. 173, no. 1, pp. 14–24, Aug. 2006.

[55] M. Backonja, E. W. Howland, J. Wang, J. Smith, M. Salinsky, and C. S. Cleeland, "Tonic changes in alpha power during immersion of the hand in cold water," *Electroencephalogr. Clin. Neurophysiol.*, vol. 79, no. 3, pp. 192–203, Sep. 1991.

[56] A. C. Chen and P. Rappelsberger, "Brain and human pain: topographic EEG amplitude and coherence mapping," *Brain Topogr.*, vol. 7, no. 2, pp. 129–140, 1994.

[57] R.-R. Nir, A. Sinai, R. Moont, E. Harari, and D. Yarnitsky, "Tonic pain and continuous EEG: Prediction of subjective pain perception by alpha-1 power during stimulation and at rest," *Clin. Neurophysiol.*, vol. 123, no. 3, pp. 605–612, Mar. 2012.

[58] C. Huishi Zhang, A. Sohrabpour, Y. Lu, and B. He, "Spectral and spatial changes of brain rhythmic activity in response to the sustained thermal pain stimulation," *Hum. Brain Mapp.*, vol. 37, no. 8, pp. 2976–2991, Aug. 2016.

[59] J. A. Hashmi *et al.*, "Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits," *Brain*, vol. 136, no. 9, pp. 2751–2768, Sep. 2013.

[60] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[61] T. Mullen *et al.*, "Real-Time Modeling and 3D Visualization of Source Dynamics and Connectivity Using Wearable EEG," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf.*, vol. 2013, pp. 2184–2187, 2013.

[62] A. Hyvärinen, "The Fixed-Point Algorithm and Maximum Likelihood Estimation for Independent Component Analysis," *Neural Process. Lett.*, vol. 10, no. 1, pp. 1–5, Aug. 1999.

[63] V. D. Calhoun, J. Liu, and T. Adalı, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *NeuroImage*, vol. 45, no. 1, Supplement 1, pp. S163–S172, Mar. 2009.

[64] L. Qin and B. He, "A wavelet-based time–frequency analysis approach for classification of motor imagery for brain–computer interface applications," *J. Neural Eng.*, vol. 2, no. 4, p. 65, 2005.

[65] M. Sifuzzaman, M. R. Islam, and M. Z. Ali, "Application of Wavelet Transform and its Advantages Compared to Fourier Transform," 2009.

[66] J. R. Quinlan, "Simplifying decision trees," *Int. J. Man-Mach. Stud.*, vol. 27, no. 3, pp. 221–234, Sep. 1987.

[67] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[68] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details," *Methods Find Exp Clin Pharmacol*, vol. 24, no. Suppl D, pp. 5–12, 2002.

[69] M. S. Hamalainen and J. Sarvas, "Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 2, pp. 165–171, Feb. 1989.

[70] B. He, T. Musha, Y. Okamoto, S. Homma, Y. Nakajima, and T. Sato, "Electric Dipole Tracing in the Brain by Means of the Boundary Element Method and Its Accuracy," *IEEE Trans. Biomed. Eng.*, vol. BME-34, no. 6, pp. 406–414, Jun. 1987.

[71] Y. Lai *et al.*, "Estimation of in vivo human brain-to-skull conductivity ratio from simultaneous extra- and intra-cranial electrical potential recordings," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 116, no. 2, pp. 456–465, Feb. 2005.

[72] T. F. Oostendorp, J. Delbeke, and D. F. Stegeman, "The conductivity of the human skull: results of in vivo and in vitro measurements," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 11, pp. 1487–1492, Nov. 2000.

[73] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clin. Neurophysiol.*, vol. 114, no. 9, pp. 1580–1593, Sep. 2003.

[74] J. E. Brown, N. Chatterjee, J. Younger, and S. Mackey, "Towards a Physiology-Based Measure of Pain: Patterns of Human Brain Activity Distinguish Painful from Non-Painful Thermal Stimulation," *PLOS ONE*, vol. 6, no. 9, p. e24124, Sep. 2011.

[75] G. Huang, P. Xiao, Y. S. Hung, G. D. Iannetti, Z. G. Zhang, and L. Hu, "A novel approach to predict subjective pain perception from single-trial laser-evoked potentials," *NeuroImage*, vol. 81, pp. 283–293, Nov. 2013.

[76] M. Prato, S. Favilla, L. Zanni, C. A. Porro, and P. Baraldi, "A regularization algorithm for decoding perceptual temporal profiles from fMRI data," *NeuroImage*, vol. 56, no. 1, pp. 258–267, May 2011.

[77] G. C. Cawley and N. L. C. Talbot, "On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation," *J. Mach. Learn. Res.*, vol. 11, no. Jul, pp. 2079–2107, 2010.

[78] R. V. Sharan and T. J. Moir, "Comparison of multiclass SVM classification techniques in an audio surveillance application under mismatched conditions," in *2014 19th International Conference on Digital Signal Processing*, 2014, pp. 83–88.

[79] Z. G. Zhang, L. Hu, Y. S. Hung, A. Mouraux, and G. D. Iannetti, "Gamma-Band Oscillations in the Primary Somatosensory Cortex—A Direct and Obligatory Correlate of Subjective Pain Intensity," *J. Neurosci.*, vol. 32, no. 22, pp. 7429–7438, May 2012.

[80] H. E. Rossiter, S. F. Worthen, C. Witton, S. D. Hall, and P. L. Furlong, "Gamma oscillatory amplitude encodes stimulus intensity in primary somatosensory cortex," *Front. Hum. Neurosci.*, vol. 7, Jul. 2013.

[81] M. Ploner, J. Gross, L. Timmermann, and A. Schnitzler, "Cortical representation of first and second pain sensation in humans," *Proc. Natl. Acad. Sci.*, vol. 99, no. 19, pp. 12444–12448, Sep. 2002.

[82] B. Kulkarni *et al.*, "Attention to pain localization and unpleasantness discriminates the functions of the medial and lateral pain systems," *Eur. J. Neurosci.*, vol. 21, no. 11, pp. 3133–3142, Jun. 2005.

[83] B. A. Vogt, "Pain and Emotion Interactions in Subregions of the Cingulate Gyrus," *Nat. Rev. Neurosci.*, vol. 6, no. 7, pp. 533–544, Jul. 2005.

[84] B. A. Vogt, S. Derbyshire, and A. K. P. Jones, "Pain Processing in Four Regions of Human Cingulate Cortex Localized with Co-registered PET and MR Imaging," *Eur. J. Neurosci.*, vol. 8, no. 7, pp. 1461–1473, Jul. 1996.

[85] I. A. Strigo, G. H. Duncan, M. Boivin, and M. C. Bushnell, "Differentiation of Visceral and Cutaneous Pain in the Human Brain," *J. Neurophysiol.*, vol. 89, no. 6, pp. 3294–3303, Jun. 2003.

[86] M.-X. Huang, D. L. Harrington, K. M. Paulson, M. P. Weisend, and R. R. Lee, "Temporal dynamics of ipsilateral and contralateral motor activity during voluntary finger movement," *Hum. Brain Mapp.*, vol. 23, no. 1, pp. 26–39, Sep. 2004.

[87] R. Näätänen and T. Picton, "The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure," *Psychophysiology*, vol. 24, no. 4, pp. 375–425, Jul. 1987.

[88] M. Liang, A. Mouraux, L. Hu, and G. D. Iannetti, "Primary sensory cortices contain distinguishable spatial patterns of activity for each sense," *Nat. Commun.*, vol. 4, p. 1979, Jun. 2013.

[89] K. D. Davis, E. Racine, and B. Collett, "Neuroethical issues related to the use of brain imaging: Can we and should we use brain imaging as a biomarker to diagnose chronic pain?:," *Pain*, vol. 153, no. 8, pp. 1555–1559, Aug. 2012.

[90] L. Heine *et al.*, "Resting State Networks and Consciousness," *Front. Psychol.*, vol. 3, Aug. 2012.

[91] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-State fMRI: A Review of Methods and Clinical Applications," *Am. J. Neuroradiol.*, vol. 34, no. 10, pp. 1866–1872, Oct. 2013.

[92] H. Ung, J. E. Brown, K. A. Johnson, J. Younger, J. Hush, and S. Mackey, "Multivariate Classification of Structural MRI Data Detects Chronic Low Back Pain," *Cereb. Cortex*, p. bhs378, Dec. 2012.

[93] S. W. Derbyshire, "Exploring the pain 'neuromatrix,'" *Curr. Rev. Pain*, vol. 4, no. 6, pp. 467–477, 2000.

[94] J. S. Feinstein *et al.*, "Preserved emotional awareness of pain in a patient with extensive bilateral damage to the insula, anterior cingulate, and amygdala," *Brain Struct. Funct.*, vol. 221, no. 3, pp. 1499–1511, Apr. 2016.

[95] A. J. Shackman, T. V. Salomons, H. A. Slagter, A. S. Fox, J. J. Winter, and R. J. Davidson, "The integration of negative affect, pain and cognitive control in the cingulate cortex," *Nat. Rev. Neurosci.*, vol. 12, no. 3, pp. 154–167, Mar. 2011.

[96] C. J. Price, "The anatomy of language: contributions from functional neuroimaging," *J. Anat.*, vol. 197, no. 3, pp. 335–359, Oct. 2000.

[97] A. Kucyi, M. Hodaie, and K. D. Davis, "Lateralization in intrinsic functional connectivity of the temporoparietal junction with salience- and attention-related brain networks," *J. Neurophysiol.*, vol. 108, no. 12, pp. 3382–3392, Sep. 2012.

[98] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. V. 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. 9673–9678, Jul. 2005.

[99] A. R. Laird *et al.*, "Behavioral Interpretations of Intrinsic Connectivity Networks," *J. Cogn. Neurosci.*, vol. 23, no. 12, pp. 4022–4037, Jun. 2011.

[100] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nat. Rev. Neurosci.*, vol. 8, no. 9, p. 700, Sep. 2007.

[101] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proc. Natl. Acad. Sci.*, vol. 100, no. 1, pp. 253–258, Jan. 2003.

[102] M. Case *et al.*, "Increased theta band EEG power in sickle cell disease patients," *J. Pain Res.*, vol. 11, pp. 67–76, Dec. 2017.

[103] M. D. Fox, R. L. Buckner, H. Liu, M. M. Chakravarty, A. M. Lozano, and A. Pascual-Leone, "Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases," *Proc. Natl. Acad. Sci.*, vol. 111, no. 41, pp. E4367–E4375, Oct. 2014.

[104] J. Cabral *et al.*, "Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations," *NeuroImage*, vol. 90, pp. 423–435, Apr. 2014.

[105] A. Mitra and M. E. Raichle, "How networks communicate: propagation patterns in spontaneous brain activity," *Phil Trans R Soc B*, vol. 371, no. 1705, p. 20150546, Oct. 2016.

[106] D. Lehmann, "Brain Electric Microstates and Cognition: The Atoms of Thought," in *Machinery of the Mind*, Birkhäuser, Boston, MA, 1990, pp. 209–224.

[107] C. M. Michel and T. Koenig, "EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review," *NeuroImage*, Dec. 2017.

[108] F. Musso, J. Brinkmeyer, A. Mobascher, T. Warbrick, and G. Winterer, "Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks," *NeuroImage*, vol. 52, no. 4, pp. 1149–1161, Oct. 2010.

[109] J. Britz, D. Van De Ville, and C. M. Michel, "BOLD correlates of EEG topography reveal rapid resting-state network dynamics," *NeuroImage*, vol. 52, no. 4, pp. 1162–1170, Oct. 2010.

[110] D. V. D. Ville, J. Britz, and C. M. Michel, "EEG microstate sequences in healthy humans at rest reveal scale-free dynamics," *Proc. Natl. Acad. Sci.*, vol. 107, no. 42, pp. 18179–18184, Oct. 2010.

[111] P. Milz, P. L. Faber, D. Lehmann, T. Koenig, K. Kochi, and R. D. Pascual-Marqui, "The functional significance of EEG microstates—Associations with modalities of thinking," *NeuroImage*, vol. 125, no. Supplement C, pp. 643–656, Jan. 2016.

[112] B. A. Seitzman, M. Abell, S. C. Bartley, M. A. Erickson, A. R. Bolbecker, and W. P. Hetrick, "Cognitive manipulation of brain electric microstates," *NeuroImage*, vol. 146, no. Supplement C, pp. 533–543, Feb. 2017.

[113] W. K. Kirchner, "Age differences in short-term retention of rapidly changing information," *J. Exp. Psychol.*, vol. 55, no. 4, pp. 352–358, Apr. 1958.

[114] J. Buhle and T. D. Wager, "Performance-dependent inhibition of pain by an executive working memory task:," *Pain*, vol. 149, no. 1, pp. 19–26, Apr. 2010.

[115] T.-P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.

[116] V. Strelets *et al.*, "Chronic schizophrenics with positive symptomatology have shortened EEG microstate durations," *Clin. Neurophysiol.*, vol. 114, no. 11, pp. 2043–2051, Nov. 2003.

[117] T. Koenig, D. Lehmann, M. C. G. Merlo, K. Kochi, D. Hell, and M. Koukkou, "A deviant EEG brain microstate in acute, neuroleptic-naive schizophrenics at rest," *Eur. Arch. Psychiatry Clin. Neurosci.*, vol. 249, no. 4, pp. 205–211, Aug. 1999.

[118] A. Khanna, A. Pascual-Leone, C. M. Michel, and F. Farzan, "Microstates in resting-state EEG: Current status and future directions," *Neurosci. Biobehav. Rev.*, vol. 49, pp. 105–113, Feb. 2015.

[119] T. Koenig *et al.*, "Millisecond by Millisecond, Year by Year: Normative EEG Microstates and Developmental Stages," *NeuroImage*, vol. 16, no. 1, pp. 41–48, May 2002.

[120] M. Gschwind *et al.*, "Fluctuations of spontaneous EEG topographies predict disease state in relapsing-remitting multiple sclerosis," *NeuroImage Clin.*, vol. 12, pp. 466–477, 2016.

[121] M. Gschwind, "Long-range dependencies make the difference—Comment on 'A stochastic model for EEG microstate sequence analysis,'" *NeuroImage*, vol. 117, pp. 449–455, Aug. 2015.

[122] C.-K. Peng *et al.*, "Long-range correlations in nucleotide sequences," *Nature*, vol. 356, no. 6365, p. 168, Mar. 1992.

[123] P. C. Ivanov *et al.*, "Multifractality in human heartbeat dynamics," *Nature*, vol. 399, no. 6735, p. 461, Jun. 1999.

[124] H. Wendt, "Contributions of Wavelet Leaders and Bootstrap to Multifractal Analysis: Images, Estimation Performance, Dependence Structure and Vanishing Moments. Confidence Intervals and Hypothesis Tests.," Theses, Ecole normale supérieure de lyon - ENS LYON, 2008.

[125] R. Ptak, "The Frontoparietal Attention Network of the Human Brain: Action, Saliency, and a Priority Map of the Environment," *The Neuroscientist*, vol. 18, no. 5, pp. 502–515, Oct. 2012.

[126] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nat. Rev. Neurosci.*, vol. 3, no. 3, p. 201, Mar. 2002.

[127] C. Rottschy *et al.*, "Modelling neural correlates of working memory: A coordinate-based meta-analysis," *NeuroImage*, vol. 60, no. 1, pp. 830–846, Mar. 2012.

[128] A. Kucyi and K. D. Davis, "The dynamic pain connectome," *Trends Neurosci.*, vol. 38, no. 2, pp. 86–95, Feb. 2015.

[129] D. Sridharan, D. J. Levitin, and V. Menon, "A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks," *Proc. Natl. Acad. Sci.*, vol. 105, no. 34, pp. 12569–12574, Aug. 2008.

[130] D. A. Seminowicz and K. D. Davis, "Pain Enhances Functional Connectivity of a Brain Network Evoked by Performance of a Cognitive Task," *J. Neurophysiol.*, vol. 97, no. 5, pp. 3651–3659, May 2007.

[131] H. Yuan, V. Zotev, R. Phillips, W. C. Drevets, and J. Bodurka, "Spatiotemporal dynamics of the brain at rest — Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks," *NeuroImage*, vol. 60, no. 4, pp. 2062–2072, May 2012.

[132] G. Pourtois, S. Delplanque, C. Michel, and P. Vuilleumier, "Beyond Conventional Event-related Brain Potential (ERP): Exploring the Time-course of Visual Emotion Processing Using Topographic and Principal Component Analyses," *Brain Topogr.*, vol. 20, no. 4, pp. 265–277, Jun. 2008.

[133] M. M. Murray, D. Brunet, and C. M. Michel, "Topographic ERP Analyses: A Step-by-Step Tutorial Review," *Brain Topogr.*, vol. 20, no. 4, pp. 249–264, Jun. 2008.

[134] T. Rukat, A. Baker, A. Quinn, and M. Woolrich, "Resting state brain networks from EEG: Hidden Markov states vs. classical microstates," *ArXiv160602344 Q-Bio Stat*, Jun. 2016.

[135] S.-C. Wang, "Artificial Neural Network," in *Interdisciplinary Computing in Java Programming*, Springer, Boston, MA, 2003, pp. 81–100.

[136] R. Hecht-nielsen, "III.3 - Theory of the Backpropagation Neural Network*," in *Neural Networks for Perception*, H. Wechsler, Ed. Academic Press, 1992, pp. 65–93.

[137] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Gated Feedback Recurrent Neural Networks," in *International Conference on Machine Learning*, 2015, pp. 2067–2075.

[138] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *ArXiv13013781 Cs*, Jan. 2013.

[139] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.