UNIVERSITY OF CALIFORNIA SAN DIEGO

Deep learning methods for Electrocorticographic data analyses

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Electrical Engineering
(Intelligent Systems, Robotics, and Control)

by

Tejaswy Pailla

Committee in charge:

Professor Vikash Gilja, Chair
Professor Virginia de Sa
Professor Kenneth Kreutz-Delgado
Professor Truong Nguyen
Professor Lawrence K. Saul

2019

The dissertation of Tejaswy Pailla is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

EPIGRAPH

The questions are always more important than the answers.

Professor Randy Pausch

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This work would not have been possible without the immense support of so many people. First and foremost, I would like to thank my advisor and mentor, Dr.Vikash Gilja for all the deep and fruitful interactions over the years. His guidance and support has kept me going through the years. I hope to embody his dedication to research, patience and cordiality in my research and personal life.

My colleagues in the Translational Neural Engineering Lab made my days in lab fun and exciting. Special thanks to Paolo Gabriel for the awesome company over the last few years. Thanks to Hannah for dragging me on runs and John, Molly for the hikes. It was a pleasure working with Werner Jiang, Anthony Au, Venkatesh Elango, Aashish Patel, Nick Rogers, Kenny Chen, Akinyinka Omigbodun, Wahab Alasfour, Daril Brown and Yuchen Wang.

I am indebted to my parents for their love and support. Thank you for being patient with me over the last few years. Special thanks to my mom for always standing by my side and my dad for trusting me with my choices. I am lucky to have the coolest in-laws and their love and support over the last two years has been amazing. My husband, Teja has been a rock. Thanks for relentlessly striving to bring out the best in me. I love you.

The chapters of this dissertation consist of published and submitted journal articles. The dissertation author is the primary author of each of these papers.

Chapter 2, in part, is a reprint of the material from material as it appears in the Proceedings of the 2016 Engineering in Medicine and Biology Conference, "ECoG data analyses to inform closed-loop BCI experiments for speech-based prosthetic applications." Pailla, Tejaswy, Werner Jiang, Benjamin Dichter, Edward F. Chang, and Vikash Gilja. The dissertation author was the primary investigator and author of this paper. Chapter 2 also contains analyses done together with Anthony Au and is unpublished.

Chapter 3, in full, is a reprint of the material as it appears in Journal of Neural Engineering (2018), "Autoencoders for learning template spectrograms in electrocorticographic signals." Pailla, Tejaswy, Kai Joshua Miller, and Vikash Gilja. The dissertation author is the primary investigator and author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. "Cross-subject learning for decoding electrocorticographic signals". Pailla, Tejaswy, Venkatesh Elango, Aashish Patel, Vikash Gilja. The dissertation author is the primary investigator and author of this paper.

VITA

2012	Summer research intern, Vanderbilt University
2013	Bachelor of Technology (Honors) in Electrical Engineering, Indian Institute of Technology-Bhubaneswar, India
2013 – 2019	Graduate Student Researcher, University of California, San Diego
2016 – 2019	Teaching Assistant, University of California, San Diego
2015	Master of Science in Electrical Engineering (Intelligent systems, Robotics & Control), University of California, San Diego
2015	Summer research intern, IBM Almaden Research Center, San Jose
2018	Software engineering intern, CliniComp Iternational, San Diego
2019	Doctor of Philosophy in Electrical Engineering (Intelligent Systems, Robotics & Control), University of California, San Diego

PUBLICATIONS

Pailla, Tejaswy and Vikash Gilja. "Cross subject transfer learning using LSTMs for ECoG based BCIs." (In Preparation)

Gabriel, Paolo Gutierrez, Kenny Chen, Abdulwahab Alasfour, **Tejaswy Pailla**, Werner Doyle, Orrin Devinsky, Daniel Friedman et al. "Neural correlates of unstructured motor behaviors." Journal of Neural Engineering (2019).

Pailla, Tejaswy, Kai Joshua Miller, and Vikash Gilja. "Autoencoders for learning template spectrograms in electrocorticographic signals." Journal of Neural Engineering (2018).

Wang, Kaiping, Udit Parekh, **Tejaswy Pailla**, Harinath Garudadri, Vikash Gilja, and Tse Nga Ng. "Stretchable dry electrodes with concentric ring geometry for enhancing spatial resolution in electrophysiology." Advanced healthcare materials 6, no. 19 (2017): 1700552.

Pailla, Tejaswy, Werner Jiang, Benjamin Dichter, Edward F. Chang, and Vikash Gilja. "ECoG data analyses to inform closed-loop BCI experiments for speech-based prosthetic applications." In Engineering in Medicine and Biology Society (EMBC), 2016. *(Selected for lecture presentation)*

Jiang, Werner, **Pailla, Tejaswy**, Benjamin Dichter, Edward F. Chang, and Vikash Gilja. "Decoding speech using the timing of neural signal modulation." In Engineering in Medicine and Biology Society (EMBC), 2016.

Ghosh A, Krishnan TR, **Pailla, Tejaswy**, Mandal A, Pradhan JK, Ranasingh S. Design and implementation of a 2-DOF PID compensation for magnetic levitation systems. ISA transactions. 2014 Jul 1;53(4):1216-22.

ABSTRACT OF THE DISSERTATION

Deep learning methods for

Electrocorticographic data analyses

by

Tejaswy Pailla

Doctor of Philosophy in Electrical Engineering

(Intelligent Systems, Robotics, and Control)

University of California San Diego, 2019

Professor Vikash Gilja, Chair

Electrocorticography (ECoG) records brain activity from the cortical surface. ECoG data analyses has led to significant advancements in neuroscientific research, particularly in two major domains: functional mapping to understand the cortical organization of human brain; and brain-machine interfaces (BMIs) that decode intent from neural data. Designing high performance BMIs is an active area of interest. A discrete BMI design primarily involves decoding specific targets from features extracted from ECoG data. Majority of ECoG based research studies use spectral features i.e. powers in specific frequency bands;

which are based on empirical observations. However, given the non-stationarity and variability of neural signals, features extracted in a data driven way could lead to more robust BMIs. In addition to efficient feature extraction and decoder training, the choice of targets presented to BMI user can greatly affect the bit-rate or throughput of the BMI.

The ability to record ECoG data in the order of days in epilepsy monitorig units (EMU); in synchrony with behavioral data through non-invasive sensors like Kinect; has resulted in deluge of large-scale, coarsely labelled datasets. Deep learning architectures are being explored to tackle this big-data problem and extract meaningful spatio-temporal patterns from these data. However, much of the existing research has been relying on architectures that found success in other domains such as computer vision and audio. A more systematic approach towards neural network architecture designs for analyzing large-scale ECoG datasets, that embeds domain knowledge from neuroscience and neurophysiology, is necessary.

The contributions of this thesis are three-fold. Firstly, I show that we can increase the throughput of a speech-based BMI by using targets that generate maximal separation in neural space. Secondly, I show that data-driven features can be learnt in an unsupervised fashion using autoencoders and are more robust compared to linear dimensionality reductions methods like PCA. These features also aid in funcional mapping by identifying functionally similar electrodes in unsupervised fashion. Lastly, I show that cross-subject model can be learnt to decode finger flexions by learning common latent spaces that map activity from different subjects onto a single latent space. By learning temporal dynamics using a recurrent neural network, from this common space, we show that we can decode continuous behvaiors from ECoG.

By analyzing the architectural design decisions on smaller, well-structured and labelled datasets, we can have a smarter approach towards developing deep learning toolkits for larger, coarsely labelled or unlabelled datasets. The methods described in this thesis will aid neuroscientists in ECoG data analysis, clinicians by providing data driven functional mapping methods, and neural engineers by providing more robust machine learning pipelines for BMI design.

Chapter 1

Introduction

Nothing in neuroscience makes sense except in the light of behavior

-Shepherd, Gordon M. Neurobiology. Oxford University Press, 1988. [1]

1.1 Background

1.1.1 Neural recording technologies

Neural recording methods provide a way to record brain activity. Invasive neural recording methods such as intracortical spike recordings, Electrocorticography (ECoG), or non-invasive recording modalities like functional near-infrared spectroscopy (fNIRS), function Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG) have various clinical applications depending on the spatial, temporal resolutions and the invasiveness of recording techniques. Figure 1.1 shows various recording modalities along with their spatial and temporal resolutions. Depending on the spatial and temporal resolution and the kind of activity measured, e.g. electrical potentials or blood oxygenation levels, these recording techniques have enabled great progress in our understanding of the functioning of brain.

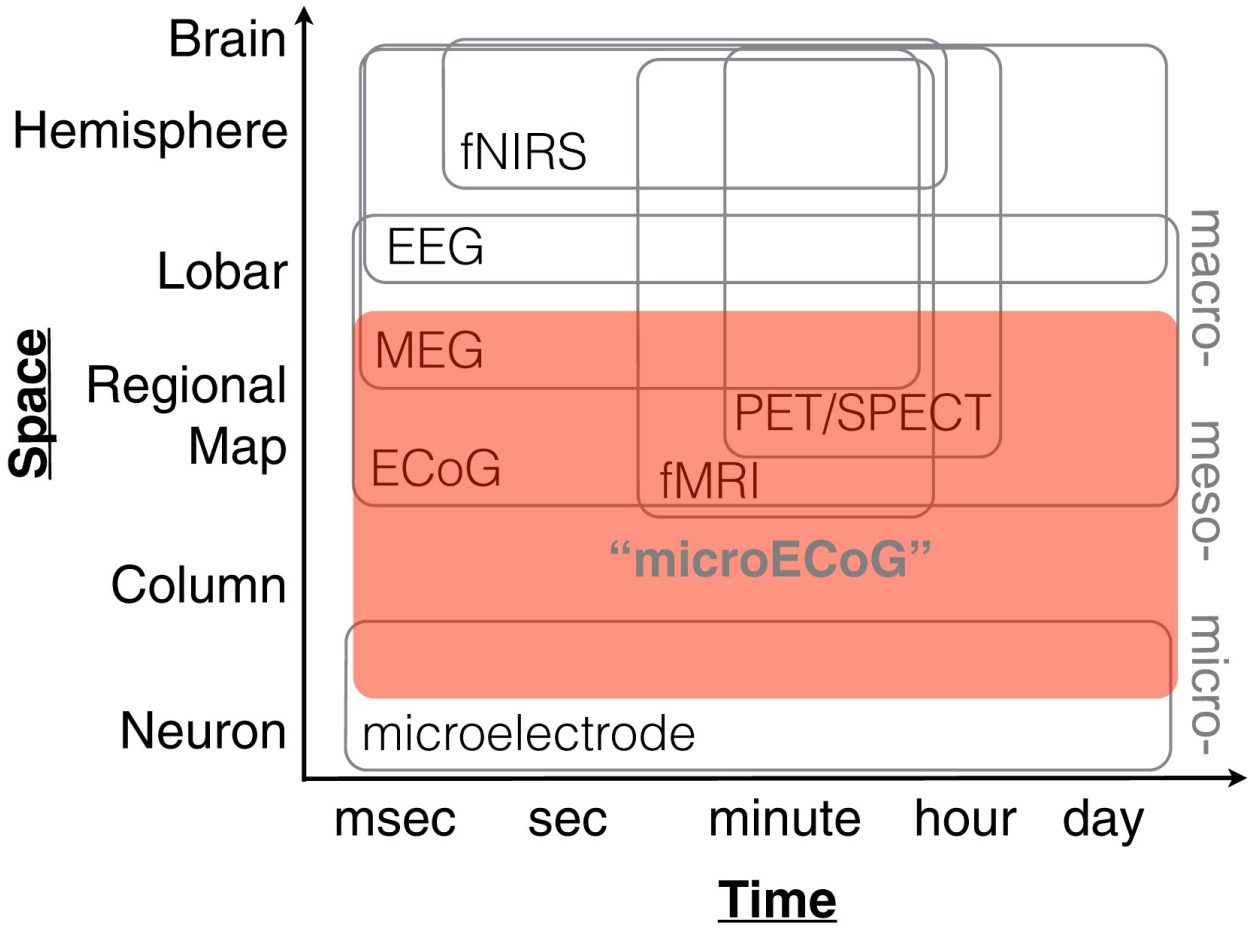

Figure 1.1: Dimensions of Spatial and Temporal Resolution in Human Neurophysiology from [2]

1.1.2 Electrocorticography

ECoG records local field potentials from the cortical surface. ECoG is primarily used to monitor cortical activity for identifying epileptic seizure foci. Individuals with chronic epilepsy undergo surgery for electrode implantation and are placed in epilepsy monitoring units (EMUs) for post surgical care and their cortical activity is monitored to localize seizure foci [3]. Figure 1.2 shows the implanted electrode grid for epilepsy monitoring. Individuals in EMUs volunteering to participate in research studies are vital contributors to ECoG research. The data collected in EMUs has largely contributed to the advancements in neuroscientific research. Two such research areas are functional mapping ; i.e. understanding the organization of brain regions and their relationships to behavior; and the development of brain-machine interfaces

(BMIs).

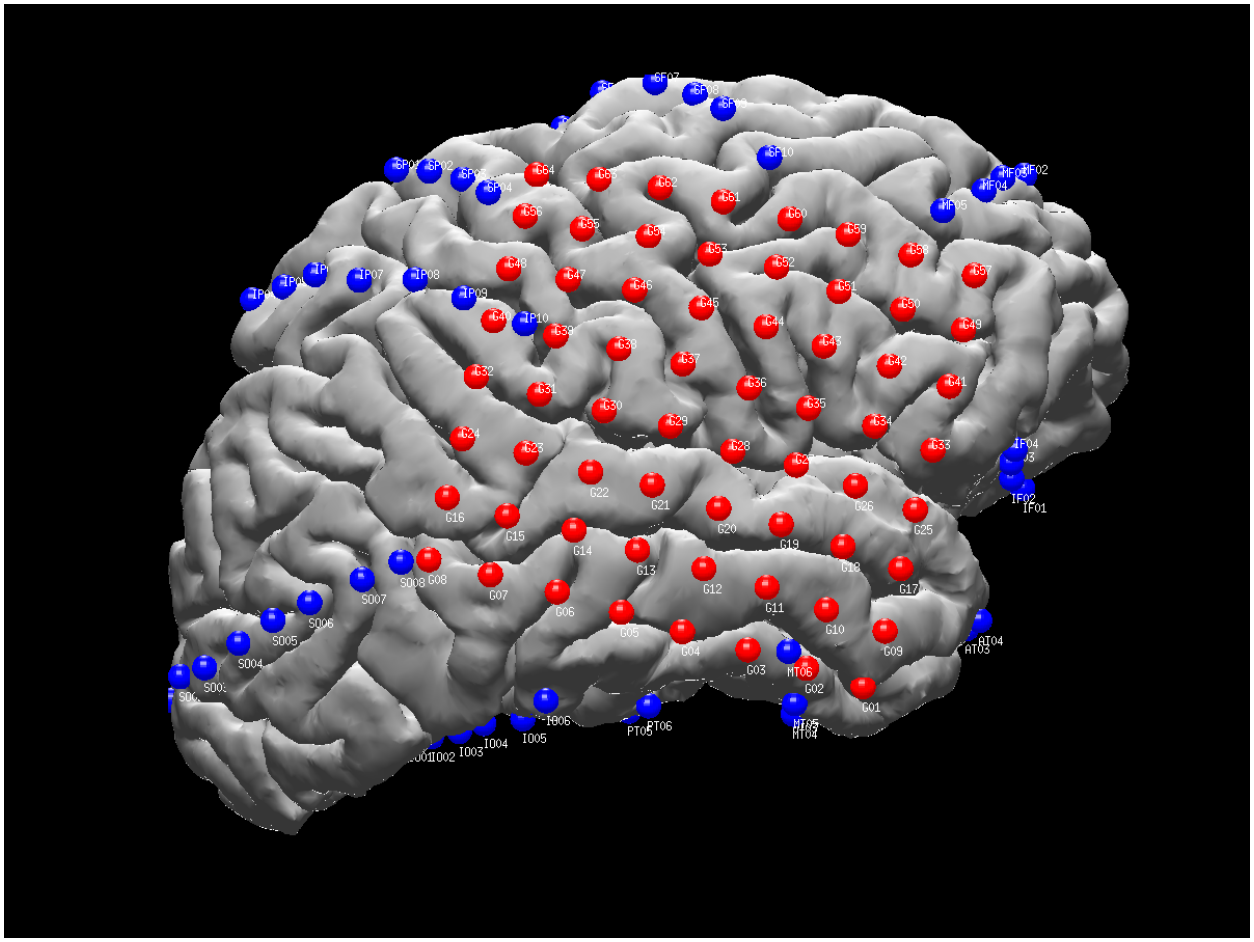

Figure 1.2: Clinical ECoG and sEEG electrodes implanted for epilespy monitoring

1.1.3 Functional mapping

Analyzing ECoG data has advanced our understanding of motor, cognitive behaviors and memory [4–6]. Offline analysis of data from structured experimental tasks wherein the subjects perform specific behavioral activity has led to advancements in our understanding of functional organization of cortical areas. For example, in [4], hand areas were mapped using spectral powers in high frequency band (76-200 Hz). In [7], human sensorimotor cortex has been mapped for speech production from the data obtained as subjects spoke a diverse set of phonemes that enabled timed movements of articulators, thereby enabling

researching to study the relationship between cortical area and the articulator.

1.1.4 Brain-machine interfaces

BMIs decode intent from neural data by modeling spatio-temporal patterns. ECoG records neural data from the cortical surface and has higher spatial and temporal resolution compared to non-invasive neural recording techniques such as EEG, fMRI. The viability of ECoG as a neural acquisition modality for prosthetic control has been an area of active research for the past two decades ([8–11]. The general pipeline of ECoG based BMI studies is shown in top part of Figure 1.4. The decoder is usually built using prior recorded training data after extensive offline analyses to perform feature extraction and model selection. [12, 13] demonstrated the use of ECoG based BMIs in paralyzed individuals with great success. In [14], the authors show the feasibility of wireless implantations of ECoG devices that can record signal and decode it to give control signal to BMIs in real time.

1.2 Motivating problem

Spectral features i.e. powers in low frequency band (Beta : 8-32 Hz) and high frequency band (High-γ : 76-150 Hz) have been traditionally dominated ECoG-based research studies [15]. Strong evidence of decrease in low frequency power and increase in high-γ power during motor movements, as shown in figure 1.3, has motivated majoritty of ECOG based BMI studies to rely on these spectral features.

Need for data driven features: There has been an increased interest in the development of high density ECoG electrodes [16]. The BRAIN initiative has propelled the development of tools and techniques to record brain activity at a single neuron level, understand neural circuitry and mapping neural activity to behaviors, perceptions, thoughts and emotions [17, 18]. The electrode diameter and inter-electrode distance, i.e. the pitch has been steadily reducing from 1 cm in clinical grade ECoG to micron scale, allowing researchers to record neural data at varying resolutions, from thousands of neurons to single neuron reso-

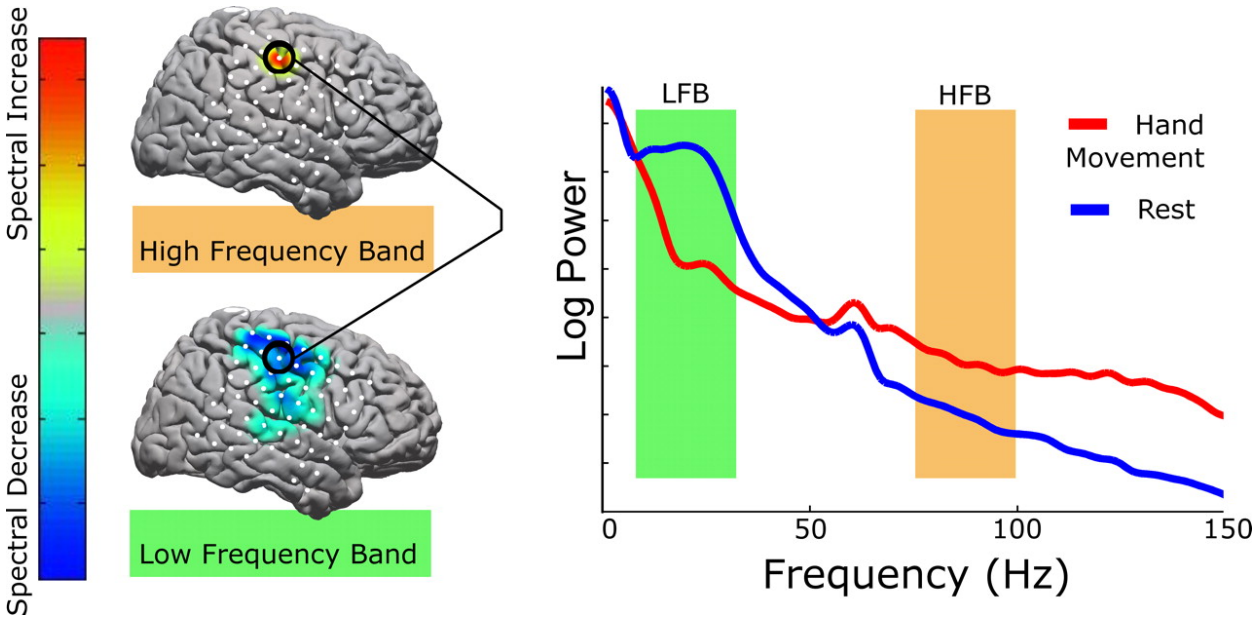

Figure 1.3: Spectral changes in cortical surface potentials during motor movement. Source [15].

lution. For example, using Neurogrid, a thin flexible, high density micro-ECoGdevice is shown to record single units in rats [19]. In [20], micro ECoG with a pitch of 400 microns has been shown to record evoked activity has been recording in humans in operating room. The ability to record ECoG at varying resolution warrants new data driven features rather than relying on empirical observations.

Cross-subject learning: Despite the moderate dataset sizes in brain imaging research, deep learning methods are being employed with considerable success in different non-invasive neuro-imaging techniques like fMRI and EEG. The standardization of fMRI and EEG data across subjects and experimental studies is also well studied, facilitating the creation of data corpuses and the development of cross-subject analysis techniques. Given the higher relative invasiveness, well labelled ECoG datasets with high trial size are scarce and inter subject variability is high as the electrode grid placement is guided by individual subject's clinical requirements. These constraints pose challenge to developing cross-subject models for BMIs.

Need for scalable, unsupervised machine learning models: Recently, the opportunity to record

long term ECoG recordings from EMUs along with behavioral data with minimally intrusive video based sensors has facilitated the creation of larger scale data corpus [21, 22]. However these datasets ae highly noisy owing to the clinical environment in which they are recorded and are largely unlabelled or coarsely labelled. In order to be able to extract meaningful spatio-temporal patterns from these data, we need to explore scalable and unsupervised machine learning methods.

In this thesis, I address the challenges of learning data driven features using unsupervised architectures and cross-subject model training. I describe tools and techniques; specifically neural network architectures; to model spatial, frequency and temporal dynamics in ECoG data. By understanding the modeling capabilities and limitations neural network architectures from smaller, well labelled datsets from structured experimental tasks, we can have a smarter approach to developing toolkits for larger datasets. The methods developed will aid neuroscientists in ECoG data analysis, clinicians by providing data driven functional mapping methods, and neural engineers by developing more robust machine learning pipelines for BMI design.

1.3 Chapter Previews

In this thesis, I explore various considerations while analyzing ECoG data, both for BMI applications and for understanding spatio-temporal structures in data for better experimental designs and aiding neuroscientific research.

In Chapter 2, I investigate possible feature extraction methods and experimental paradigms for a speech-based BCI system. We find that the exact method to estimate spectral powers doesn't have any significant effect. We show that improvement in bit rate can be achieved by a task design that allows more articulator movements in a given time.

In chapter 3, we propose an autoencoder based approach for summarizing ECoG data with "template spectrograms", i.e. informative time-frequency (t-f) patterns, and demonstrate their efficacy in two

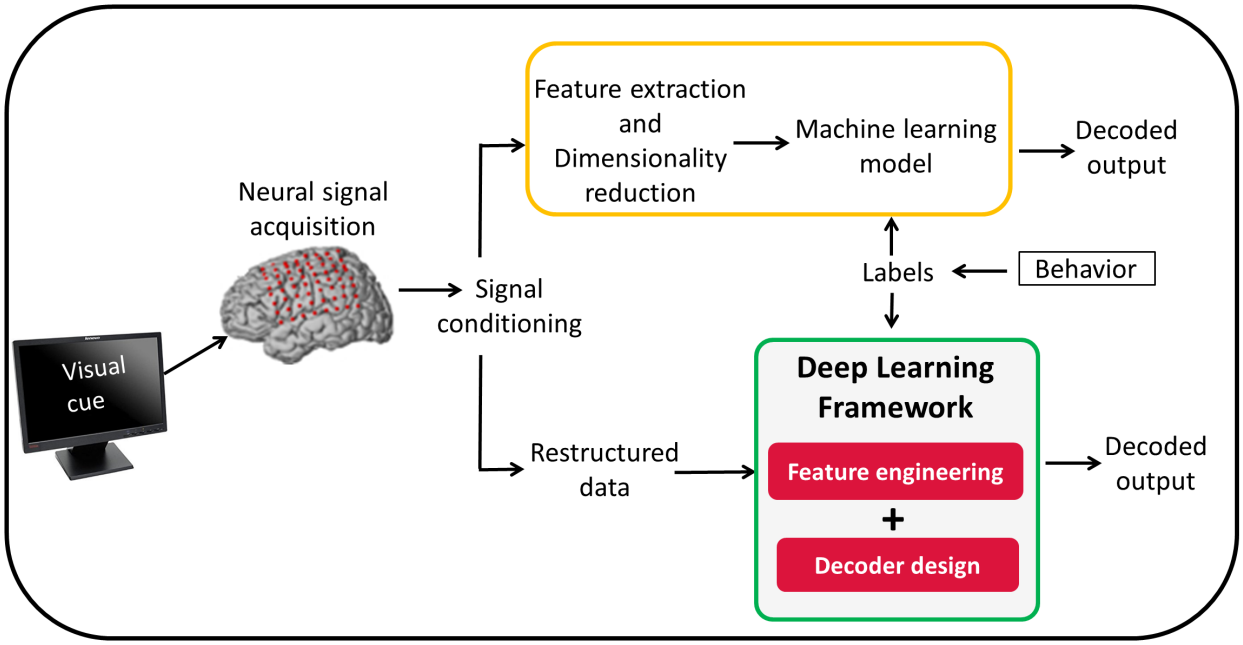

Figure 1.4: Schematic of proposed BMI decoder pipeline

contexts: brain-computer interfaces (BCIs) and functional brain mapping.

In chapter 4, we show that neural network architectures can be trained to learn common latent spaces across subjects, thereby enabling cross-subject learning for ECoG decoding.

In chapter 5, neural networks are used to find patterns in large-scale unstructured data as a preliminary demonstration of future directions for this thesis work.

References

[1] Gordon M Shepherd. *Neurobiology*. Oxford University Press, 1988.

[2] Terrence J Sejnowski, Patricia S Churchland, and J Anthony Movshon. Putting big data to good use in neuroscience. *Nature neuroscience*, 17(11):1440, 2014.

[3] Ivan Osorio, Mark G Frei, Jon Giftakis, Tom Peters, Jeff Ingram, Mary Turnbull, Michele Herzog, Mark T Rise, Scott Schaffner, Richard A Wennberg, et al. Performance reassessment of a real-time seizure-detection algorithm on long ecog series. *Epilepsia*, 43(12):1522–1535, 2002.

[4] Kai J Miller, Pradeep Shenoy, John W Miller, Rajesh PN Rao, Jeffrey G Ojemann, et al. Real-time functional brain mapping using electrocorticography. *Neuroimage*, 37(2):504–507, 2007.

[5] Andreas K Engel, Christian KE Moll, Itzhak Fried, and George A Ojemann. Invasive recordings from the human brain: clinical insights and beyond. *Nature Reviews Neuroscience*, 6(1):35, 2005.

[6] Joshua Jacobs and Michael J Kahana. Direct brain recordings fuel advances in cognitive electrophysiology. *Trends in cognitive sciences*, 14(4):162–171, 2010.

[7] Kristofer E Bouchard, Nima Mesgarani, Keith Johnson, and Edward F Chang. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441):327–332, 2013.

[8] Eric C Leuthardt, Gerwin Schalk, Jonathan R Wolpaw, Jeffrey G Ojemann, and Daniel W Moran. A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, 1(2):63, 2004.

[9] Philip R Kennedy, M Todd Kirby, Melody M Moore, Brandon King, and Adon Mallory. Computer control using human intracortical local field potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(3):339–344, 2004.

[10] Takufumi Yanagisawa, Masayuki Hirata, Youichi Saitoh, Haruhiko Kishima, Kojiro Matsushita, Tetsu Goto, Ryohei Fukuma, Hiroshi Yokoi, Yukiyasu Kamitani, and Toshiki Yoshimine. Electrocorticographic control of a prosthetic arm in paralyzed patients. *Annals of neurology*, 71(3):353–361, 2012.

[11] Guy Hotson, David P McMullen, Matthew S Fifer, Matthew S Johannes, Kapil D Katyal, Matthew P Para, Robert Armiger, William S Anderson, Nitish V Thakor, Brock A Wester, et al. Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject. *Journal of neural engineering*, 13(2):026017, 2016.

[12] Wei Wang, Jennifer L Collinger, Alan D Degenhart, Elizabeth C Tyler-Kabara, Andrew B Schwartz, Daniel W Moran, Douglas J Weber, Brian Wodlinger, Ramana K Vinjamuri, Robin C Ashmore, et al. An electrocorticographic brain interface in an individual with tetraplegia. *PloS one*, 8(2):e55344, 2013.

[13] Alan D Degenhart, Shivayogi V Hiremath, Ying Yang, Stephen Foldes, Jennifer L Collinger, Michael Boninger, Elizabeth C Tyler-Kabara, and Wei Wang. Remapping cortical modulation for electrocorticographic brain–computer interfaces: a somatotopy-based approach in individuals with upper-limb paralysis. *Journal of neural engineering*, 15(2):026021, 2018.

[14] Kojiro Matsushita, Masayuki Hirata, Takafumi Suzuki, Hiroshi Ando, Yuki Ota, Fumihiro Sato, Shyne Morris, Takeshi Yoshida, Hidetoshi Matsuki, and Toshiki Yoshimine. Development of an implantable

wireless ecog 128ch recording device for clinical brain machine interface. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1867–1870. IEEE, 2013.

[15] Kai J Miller, Eric C Leuthardt, Gerwin Schalk, Rajesh PN Rao, Nicholas R Anderson, Daniel W Moran, John W Miller, and Jeffrey G Ojemann. Spectral changes in cortical surface potentials during motor movement. *The Journal of neuroscience*, 27(9):2424–2432, 2007.

[16] Edward F Chang. Towards large-scale, human-based, mesoscopic neurotechnologies. *Neuron*, 86(1):68–78, 2015.

[17] Thomas R Insel, Story C Landis, and Francis S Collins. The nih brain initiative. *Science*, 340(6133):687–688, 2013.

[18] Cornelia Bargmann, William Newsome, A Anderson, E Brown, K Deisseroth, J Donoghue, P MacLeish, E Marder, R Normann, J Sanes, et al. Brain 2025: a scientific vision. *Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Working Group Report to the Advisory Committee to the Director, NIH*, 2014.

[19] Dion Khodagholy, Jennifer N Gelinas, Thomas Thesen, Werner Doyle, Orrin Devinsky, George G Malliaras, and György Buzsáki. Neurogrid: recording action potentials from the surface of the brain. *Nature neuroscience*, 18(2):310, 2015.

[20] Mehran Ganji, Erik Kaestner, John Hermiz, Nick Rogers, Atsunori Tanaka, Daniel Cleary, Sang Heon Lee, Jospeh Snider, Milan Halgren, Garth Rees Cosgrove, et al. Development and translation of pedot: Pss microelectrodes for intraoperative monitoring. *Advanced Functional Materials*, 28(12):1700232, 2018.

[21] Paolo Gabriel, Werner K Doyle, Orrin Devinsky, Daniel Friedman, Thomas Thesen, and Vikash Gilja. Neural correlates to automatic behavior estimations from rgb-d video in epilepsy unit. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3402–3405. IEEE, 2016.

[22] Nancy Wang, Ali Farhadi, Rajesh Rao, and Bingni Brunton. Ajile movement prediction: Multimodal deep learning for natural human neural recordings and video, 2018.

Chapter 2

ECoG data analyses using traditional methods for speech-based prosthetic applications

2.1 Introduction

Brain Computer Interfaces (BCIs) aid communication in paralyzed individuals by translating neural signals into control signals for prosthetic devices [1]. Clinical trials of intracortical BCIs demonstrate the potential for neural prostheses to partially restore lost motor and communication function in individuals with tetraplegia [2, 3]. In recent years, electrocorticographic (ECoG) signals have been investigated for prosthetic control [4, 5]. ECoG signals have higher spatio-temporal resolution and signal-to-noise ratio compared to electroencephalographic (EEG) signals [6], which may translate to system performance advantages .

The first demonstrations of ECoG-based BCIs used actual or imagined motor movements to control a one-dimensional computer cursor [4]. Since then, several studies have used ECoG signals to decode motor actions or imagery [7, 8] to control cursor movements. ECoG signals from sensory-motor cortices can also be used to decode speech elements. Given the high bit rate of natural human speech, successful speech

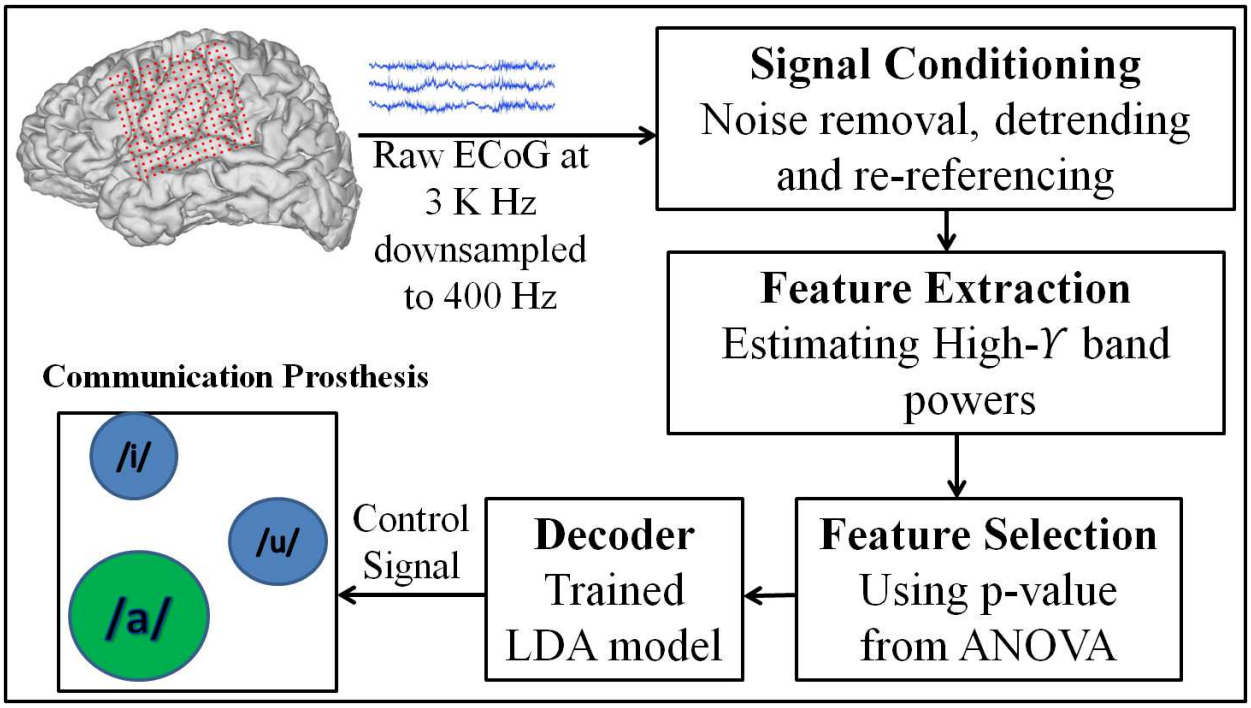

Figure 2.1: Overview of the processing pipeline

decoding could result in BCIs with high information transfer rates. Recent ECoG-based speech prosthesis studies have used whole words or combinations of consonants and vowels to decode speech elements [9–15].

Fig. 2.1 shows an overview of a BCI system. Features extracted from processed neural signals are decoded to guide a prosthetic device. Neural activity, corresponding to specific behaviors, should be well separated in the feature space to enable decoding with minimal error. Since ECoG captures activity of neural ensembles, an optimal ECoG-based BCI design paradigm would be the one that prompts realizable behavioral changes that originate from consistent and well distinguishable neural sources.

Consider the design of a speech-based BCI system that decodes individual phonemes. Vocalization of a specific phoneme requires coordinated movement of multiple articulators. The spatio-temporal dynamics of activity in sensory-motor cortex (SMC) during speech are complex, given the precise control of multiple degrees of freedom of each articulator. Correlated neural activity corresponding to articulator movements start well before speech onset and continues through vocalization [16]. Knowledge of the sepa-

rability of neural features corresponding to different behaviors can guide BCI design choices. In this study, we analyze data collected while subjects vocalize three cardinal vowel-phonemes. We analyze how behavioral paradigms affect decoding accuracy and discuss how our observations inform potential BCI design strategies. We also discuss how our findings relate to the current understanding of functional organization of SMC.

2.2 Datasets

2.2.1 Subjects

Three native English speakers; EC56, EC61, and EC69 (1 male and 2 female respectively); underwent chronic implantation of a high-density, subdural ECoG array (over the right hemisphere (EC56,61) or left hemisphere(EC69)) as part of their clinical treatment of epilepsy. Subjects gave their written informed consent before the day of surgery. The experimental protocol was approved by Human Research Protection Program at the University of California, San Francisco. All subjects had self-reported normal hearing and underwent neuropsychological language testing (including the Boston Naming and verbal fluency tests) and were found to be normal. Subjects EC56 and EC61 were implanted with 256 electrode arrays and subject EC69 with 296 electrodes. For all subjects we restricted our analysis to the 256 electrode grid covering SMC and superior temporal gyrus.

2.2.2 Recordings

Cortical local field potentials were sampled at 3052 Hz from the ECoG grid with a multi-channel amplifier optically connected to a digital signal processor (Tucker-Davis Technologies, Alachua, FL). Audio data from the subjects, in synchrony with the ECoG data, was recorded with a microphone, amplified, and digitized.

2.2.3 Experimental Paradigm

Subjects performed a call-and-response task where each subject vocalized one of three cardinal vowel-phonemes ([a], [u], [i]) at different pitch levels in their respective vocal ranges, in response to the pattern of vocalization produced by the experimenter. Fig. 2.2 shows the audio recording of one such cycle with five utterances. The duration of each utterance was 1013±283 ms, 628±49 ms and 495±34 ms for subjects EC56, 61 and 69 respectively. Each subject vocalized 6–7 such cycles for each phoneme. Baseline data, when the subject was at rest, was collected at the end of sessions.

2.3 Exploring feature extraction methods

Brain Computer Interfaces (BCIs) aid communication for paralyzed individuals by translating neural signals into control signals for a prosthetic device [1]. In recent years, electrocorticography (ECoG) based BCIs have been investigated [4, 17] as an alternative to intracortical based BCIs [18, 19]. ECoG signals have higher spatial resolution and signal-to-noise ratio (SNR) compared to electroencephalography (EEG) signals, enabling higher performance control.

Estimated band power is an effective and commonly employed signal feature used by ECoG based BCIs. For example, in [20] spectral changes in ECoG signals were analyzed during motor movements. A consistent decrease of power in low-frequency band (8–32 Hz) and increase in high-frequency band (76–100

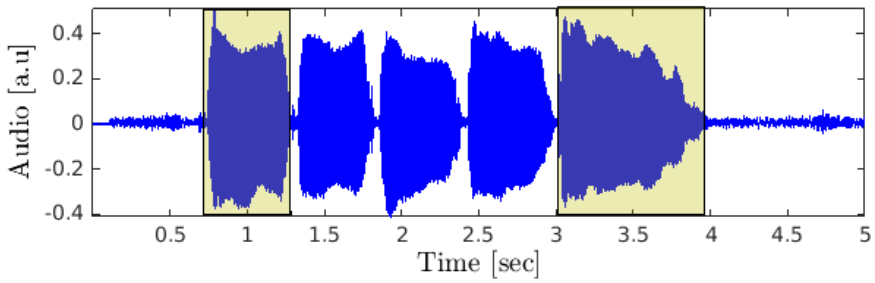

Figure 2.2: Audio recording showing a phoneme vocalization cycle where the first and last utterances (highlighted) belong to the same pitch level

Hz) across relevant electrodes was observed. In [21], band powers in low (11–40 Hz) and high(71–100 Hz) frequency ranges were used to reliably classify motor movements. Powers in higher frequency range (76–200 Hz) were found to be useful correlates of local cortical function [22]. Spectral powers in these frequency ranges prove to be reliable features for ECoG classification tasks [23].

Two common estimates of band specific powers include band-specific amplitude modulation [24] and time-varying analytic amplitude of a band-pass filtered signal obtained using Hilbert transform [25]. More recently, wavelet based spectral features were shown to give robust control features [26]. Features from wavelet transform coefficients are shown to reliably classify motor imagery signals [27].

While numerous methods of extracting spectral features have been proposed and used, power efficient computation is crucial. ECoG recording systems are progressing towards higher channel count and battery powered in vivo wireless transceivers [28, 29]. Implants with fully integrated on-board, under-the-scalp real-time signal processing will require computational choices that optimize the tradeoff between system performance and power consumption. As feature extraction operates on the broadband recorded signal and compresses these signals for follow on processing stages, it typically dominates BCI computation.

We assess the effect of commonly employed band power extraction techniques on ECoG classification performance. We test our methods on the speech vocalization dataset.

2.3.1 Methods for computing power estiamtes

Raw ECoG recorded at 3KHz is downsampled to 400 Hz before further analysis. Figure 2.6 outlines the steps of our analysis.

2.3.2 Pre-processing

We first visually and quantitatively inspected the time series data for artifacts. Electrodes 257–296 in third subject were identified as bad channels and excluded from all subsequent analysis. Electromagnetic

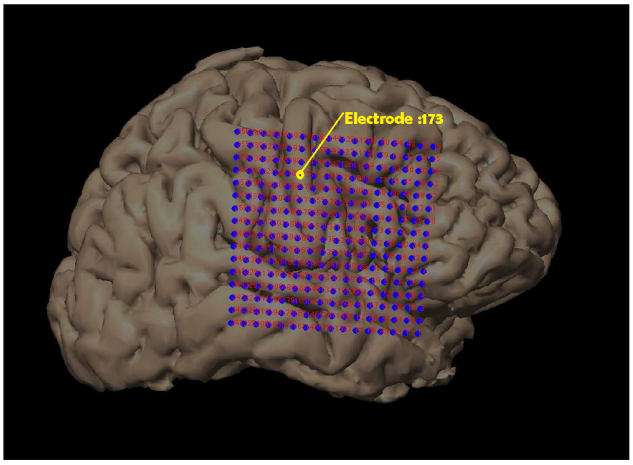

Figure 2.3: Electrode localization map of second subject (EC61) showing electrode 173 in the motor cortex.

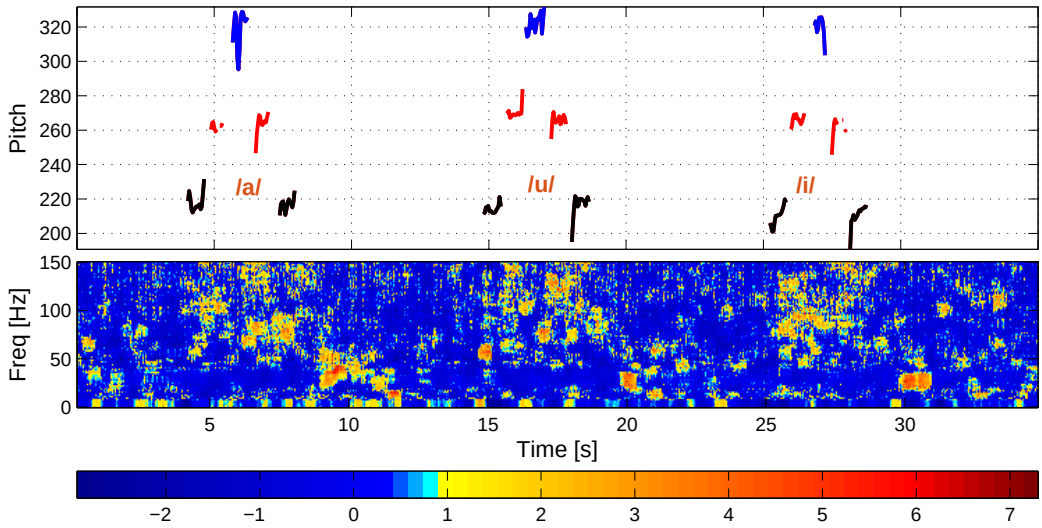

Figure 2.4: Pitch Trajectory and spectrogram for an electrode

noise common to channels was removed by common average referencing (CAR) 16-channel blocks that shared the same cable. Line noise (at 60 Hz and its harmonics) was removed using notch filters and CAR was once again performed on all the channels to remove common noise. Data is then z-scored relative to baseline and locally detrended to remove linear trending effects.

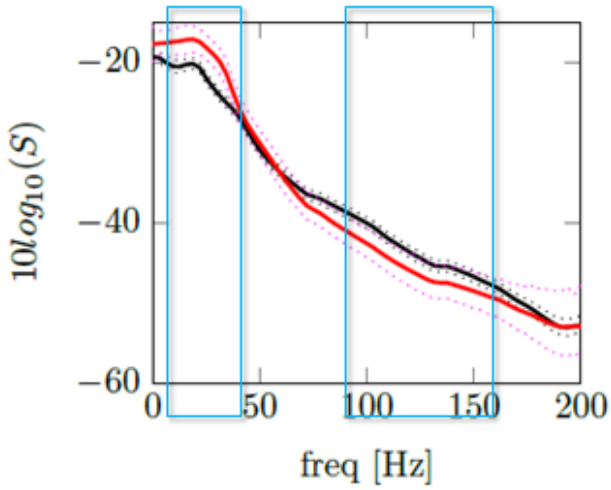

Figure 2.5: Red: Silent, Blue: Voc. The dotted lines show 95% confidence bounds derived from jackknife estimate of the standard error for each spectra

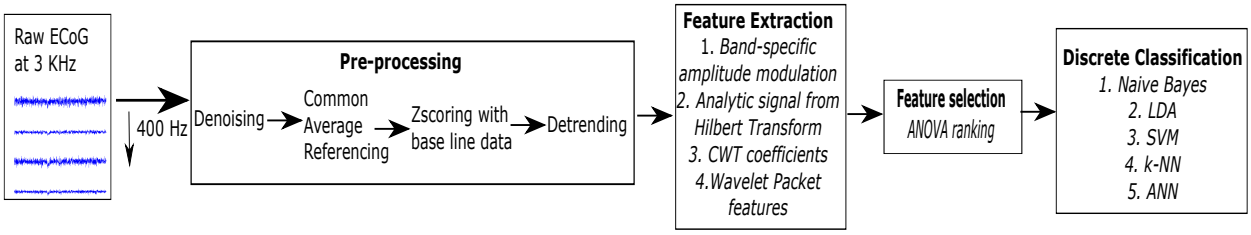

Figure 2.6: Block Diagram showing steps in analysis

2.3.3 Feature Extraction

In this work, we compute spectral features, i.e. power estimates of a window of data in two frequency ranges: low frequency band(LFB) i.e. 12.5–25 Hz, high frequency band (HFB) i.e. 75–100, using four different methods described below (M1 through M4) and evaluate their performance. The choice of frequency ranges for the two spectral bands is based on the literature and is supported by spectrograms of these data (figure 2.5). The choice of exact bounds was limited by implementation of dyadic discrete wavelet transform (DWT).

A vocalization (Voc) trial is defined as the time window from vocalization onset to offset and a

silent trial is defined as the time window from Voc offset to onset, i.e. the time between two vocalizations. Trials with less than 400 ms duration were discarded. A 300 ms time window (**W**) starting from trial onset is considered for feature extraction.

M1- Band-specific amplitude modulation(AM)

Following Sanchez et al.'s method of band-specific AM, we band-pass filtered the signal in appropriate frequency bands and mean of squared voltage in **W** was computed for feature. Let $v(t)$ be the band-pass filtered voltage signal. Spectral feature for a window of width Δt is calculated as

$$f = \frac{1}{\Delta t} \sum_{t=0}^{\Delta t} (v^2(t)) \tag{2.1}$$

M2- Analytic signal from Hilbert transform

Similar to the method followed in [16], we band pass filtered the signal using 4 Gaussian filters with logarithmically increasing center frequencies and semi-logarithmically increasing band-widths in each frequency range. We computed the analytic amplitude of these filtered signals using Hilbert transform and computed average amplitude. For each frequency band , mean of squared amplitude in **W** was taken for feature.

M3- Continuous Wavelet Transform (CWT) Coefficients

In [27], averages and standard deviations of CWT coefficients are used as features. We followed a similar approach. CWT was performed with 'symlet' mother wavelet and mean of square of coefficients in each frequency range were taken as approximation of spectral powers. Means of these power estimates in **W** were taken as features. Symlet wavelets are chosen because of their similarities with the shape of neurophysiological signals and their suitability for event-related desynchronization(ERD) analysis. [30].

M4-Wavelet Packet Features

In [30], discrete wavelet transform (DWT) based features were used. ECoG signal was first decomposed into wavelet components of specific frequency ranges by wavelet-packet analysis. Full sampled wavelet-filtered signals were reconstructed from wavelet packet components in relevant frequency ranges. We performed a 5 level DWT with the 'symlet' mother wavelet and the signal was reconstructed from appropriate DWT coefficients corresponding to LFB and HFB. Mean of squared amplitude of reconstructed signals in **W** was taken for feature.

2.3.4 Feature Selection

The construction of ECoG based classifiers in this context is subject to the "curse of dimensionality," i.e. we generally have limited training examples relative to the high dimensionality of feature space. The dimensionality of feature spaces from the above methods is number of channels $\times$ 2, (i.e. 512 for each subject).

We performed a balanced one-way ANOVA to extract relevant features. We ranked the dimensions based on p–value from the ANOVA test and used the first 15 dimensions. Note that this number was not optimized for classification performance and was only chosen to maintain reasonable dimensionality to sample size ratio.

2.3.5 Discrete Classification

We used features from the above four methods to classify ECoG data using five commonly used classification algorithms: Naive Bayes (NB), Linear Discriminant Analysis(LDA), Support Vector Machine (SVM) with a linear kernel, k-Nearest Neighbors (k-NN) with Euclidean distance as distance metric, feed-forward Artificial Neural Network (ANN) with a hidden layer of size 15. For multi-class SVM classification, we used a "One vs all" approach [31, 32]. While NB and LDA are generative models, SVM, k-NN, ANN

Table 2.1: MA and SE are mean accuracy and standard error over 5 Cross-Validation folds, M and MLA are feature extraction Method (described in section 2.3.3) and machine learning algorithm (described in section 2.3.5) that gave the best performance. Chance-level accuracies are listed in row 1.

	Voc-Sil (50%)				Pitch Level(33.33%)				Phoneme Type(33.33%)			
Subject ID	MA	SE	M	MLA	MA	SE	M	MLA	MA	SE	M	MLA
EC56	98.13	0.80	1	ANN	65.33	6.90	2	LDA	88.00	2.23	3	ANN
EC61	92.75	1.52	2	LDA	66.67	2.72	2	LDA	80.00	3.37	1	LDA
EC69	90.53	1.26	3	LDA	69.00	2.61	2	SVM	74.00	3.85	3	NB

are discriminative models. All these classification algorithms assume different models to classify data [33] and serve as good test cases for our extracted features.

2.3.6 Results comparing different power estimates

We perform 5 fold cross-validation (CV) and report mean accuracy and standard error for each classification in figure 2.7. Table 2.1 lists the best performance results for each subject and classification category. Upon referring to electrode localization maps, the top 15 dimensions from 2.10 were found to correspond to electrodes in sensorimotor cortex thereby validating our feature selection method.

Classification task largely influences the choice of feature extraction method. To distinguish between presence and absence of motor activity (eg. Voc-Sil), a crude estimate of power would suffice. Low accuracy for pitch-level classification may be attributed to the underlying complex physiology [34]. To distinguish between complex articulator movements (e.g. Phoneme type/pitch-level), a more sophisticated spectral power estimate with cleaner resolution in time and frequency and a model that better captures the time-varying dynamics of neural signal will yield better classification results.

In [35], fundamental equivalence of different spectral analysis methods namely Fourier, Hilbert and wavelet transform has been detailed. When different feature extraction methods are matched in parameter choices like window length or bandwidth, they essentially yield similar spectral power estimates. Com-

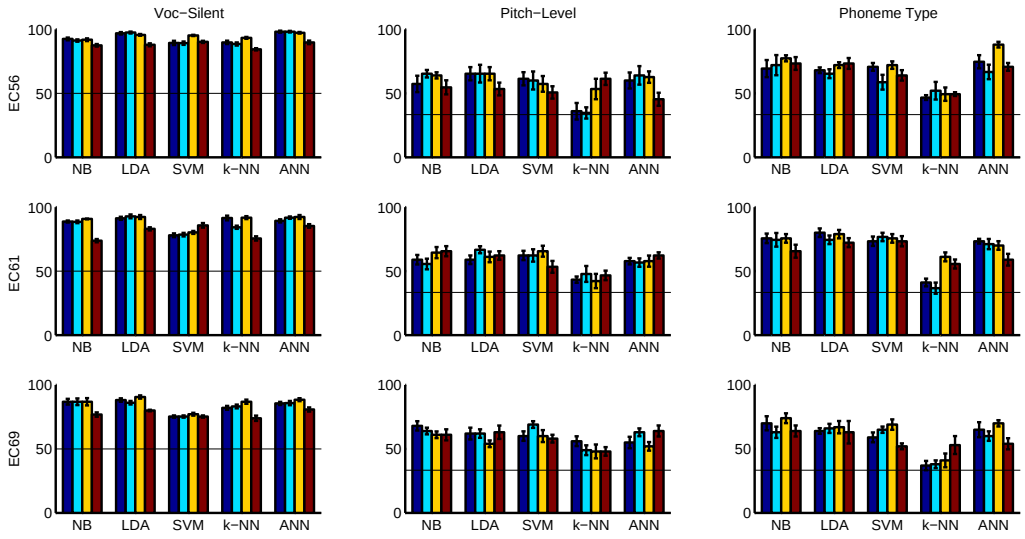

Figure 2.7: Discrete classification results. Bars represent mean accuracy over 5 CV folds for the four different feature types described in section 2.3.4. Blue: M1, Cyan: M2, Yellow: M3, Brown: M4. Errorbars represent standard error. The black horizontal line in each plot represents the chance-level accuracy (50% for Voc-Silent classification and 33.33% for Pitch Level and Phoneme classification)

putational efficiency then becomes the decisive factor in choice of feature extraction method. Of all the methods described in 2.3.3, M1 has the least computational complexity and has performance on par with more computationally intensive methods like M2 and M3. The similarity in feature spaces across all the methods can be visualized by figure 2.8. Therefore in the interests of computational costs (latency, power), M1 can be used as a robust feature extraction method for BCI applications.

2.4 Methods to inform closed-loop BCI experiments for speech-based prosthetic applications

Raw ECoG was downsampled to 400 Hz; using an anti-aliasing, linear-phase, low-pass FIR filter; before further analysis. Fig. 2.1 outlines the steps of our analysis.

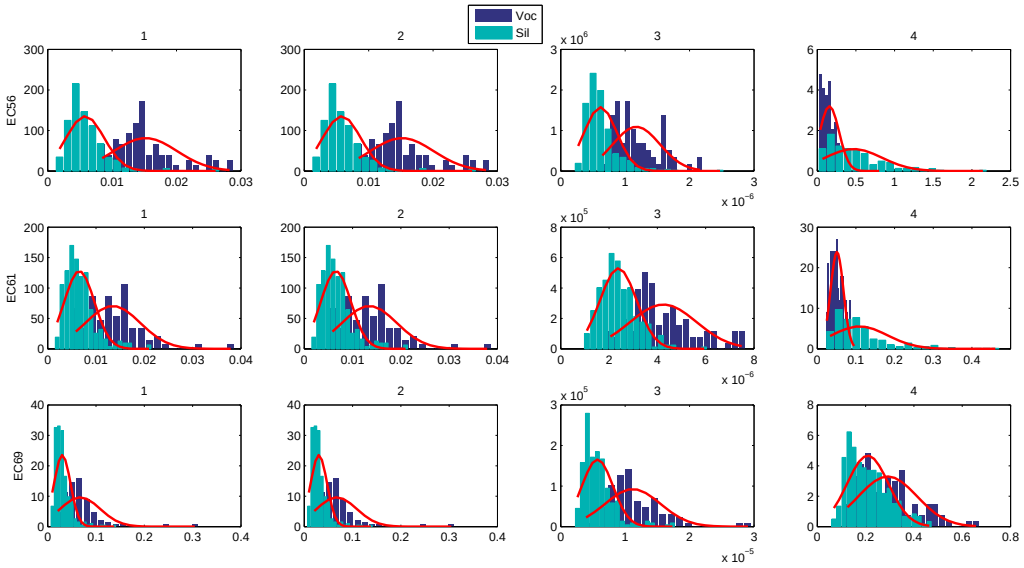

Figure 2.8: Normalized histograms showing Voc-Sil feature distributions of top ranked channel for each method across subjects .

2.4.1 Preprocessing

Electromagnetic noise common to channels was removed by common average referencing (CAR) 16-channel blocks that shared the same cable. Data was then locally detrended, using a 2 s sliding window with an overlap of 0.01 s, to remove linear trends caused by DC offsets. Line noise (at 60,120,180 Hz) was removed using a multi-taper decomposition approach [36, 37] and the resulting time series was high-pass filtered at a 1 Hz cutoff frequency. Channels with high kurtosis value ($>3\sigma$ compared to the average kurtosis of all channels combined) were identified as "bad channels" and were excluded from further analysis. CAR was once again performed on the remaining channels to remove common activity.

2.4.2 Feature Extraction and Model Selection

For our primary analysis, we considered the first and last utterances (shown in Fig. 2.2) from each vocalization cycle. These utterances are in the same pitch range, thus avoiding potential behavioral confounds with phoneme type and pitch range. For each utterance, a time window of width W ms centered

around an offset d ms from the vocalization onset was considered for feature extraction. We used band powers, from SMC electrodes, in high–γ range (75–150 Hz) to extract features [20] . To estimate the band powers, the pre-processed time series data from each channel was band-pass filtered to obtain high–γ amplitude. Squared high–γ amplitude was z-scored with baseline data and integrated in n nonoverlapping bins, each of width $\frac{W}{n}$ ms, to estimate band powers.

A regularized linear discriminant analysis (R-LDA) model with uniform prior and regularization parameter of 0.5 was used for classification. The dimensionality of the feature space is the number of channels in SMC $\times$ n and is much larger than the available number of trials. To overcome the "curse of dimensionality" and select the best features, we performed a balanced one-way ANOVA across 3 phoneme groups (using trials from both first and last utterances) and ranked features based on their p–value. Features with $p < 0.05$ were retained and used for further analysis. Optimal choice for parameters W, d, n was found via a grid search (TABLE 2.2) by maximizing the leave-one-out (LOO) cross-validation (CV) accuracy of classification of the three phonemes. Hence, no test data was used in training the decoder. We ensured that we only considered the activity during preparation and utterance of each phoneme in the grid search.

Table 2.2: Optimal parameters from grid search

Parameter	EC56	EC61	EC69
Offset in ms (d)	200	-50	200
Window width in ms (W)	400	400	300
Number of bins (n)	1	1	1

2.4.3 Decoding

When designing a speech-based communication system, the relative timing and sequencing of attempted phoneme production during training data collection and real-time use may have system perfor-

mance consequences. For example, one possible design could allow the individual to vocalize discrete phonemes, in random order, at well spaced intervals. This would require the subject to reposition his/her articulators for every utterance. Another design strategy could cue for repeated vocalization of the same phoneme. Our hypothesis is that the first paradigm that allows for more articulator movement will also generate more easily distinguishable spatio-temporal neural dynamics. Consistency of these dynamics across trials is also important to train a stable decoder. To test this, we trained the R-LDA model, using trials from two utterance types (shown in Fig. 2.2) in each vocalization cycle. The first utterance corresponds to trials that would be generated by the first design. The last utterance is a repeated vocalization of the same phoneme and simulates the trials that would be generated by the second design. We performed LOO CV to get an estimate of decoding performance. In addition, we used the R-LDA model trained on trials from one utterance type to test the trials from other utterance type to see how well the model generalizes across utterance types. The analyses are focused on measuring relative performance between decoders in two different paradigms. It should be noted that we paired the first and last utterances from the same vocalization cycle (that is left out in the LOO CV fold) for testing, thus avoiding the use of correlated neural features in training and testing sets.

2.5 Results and Discussion

We observed that for all three subjects, the channels selected by ANOVA are located in areas of SMC related to speech articulators (i.e. the lips, jaw, tongue and larynx). Fig. 2.9 shows the selected electrode locations for EC56. Although it is difficult to locate the spatial organization for individual articulators, our findings concur with the general consensus [38] of somatotopy in ventral SMC.

Fig. 2.10 shows the first three dimensions of features projected onto the LDA subspace for each utterance type. According to the LDA model, these are the linear dimensions in which the phoneme specific neural activity is most separable. We noted that in this view phoneme separability is higher in the trials

of first utterance type. Fig. 2.11 shows that first utterances are more distinguishable than last utterances suggesting that the neural activity during the first utterances is better distinguishable and more predictive of phonemes than the activity during repeated vocalizations. These results suggest that a BCI which allows more change in articulator positions (e.g. in our case, from rest to phoneme vocalization) might lead to higher bit rates than a task which doesn't encourage change in articulator positions (e.g. repeated vocalizations).

Certain phoneme pairs are more separated in neural space than others. To maximize bit rate one might choose to use only those phoneme pairs which have high information transfer rates in the BCI design. The confusion matrices (Fig. 2.12) for subjects EC61 and EC69 show confusion between phoneme pairs. To see which phoneme pairs are most separated in neural space, we performed pairwise binary classification on trials from first utterances. Fig. 2.13 shows confusion matrices from LOO CV for each phoneme pair. We calculated the Information Transfer Capacity (ITC) in bits per trial (bpt), using the Blahut-Arimoto algorithm [39]. The ITC for EC69 when phoneme pair '/a/-/u/' is used is higher than the case when all the

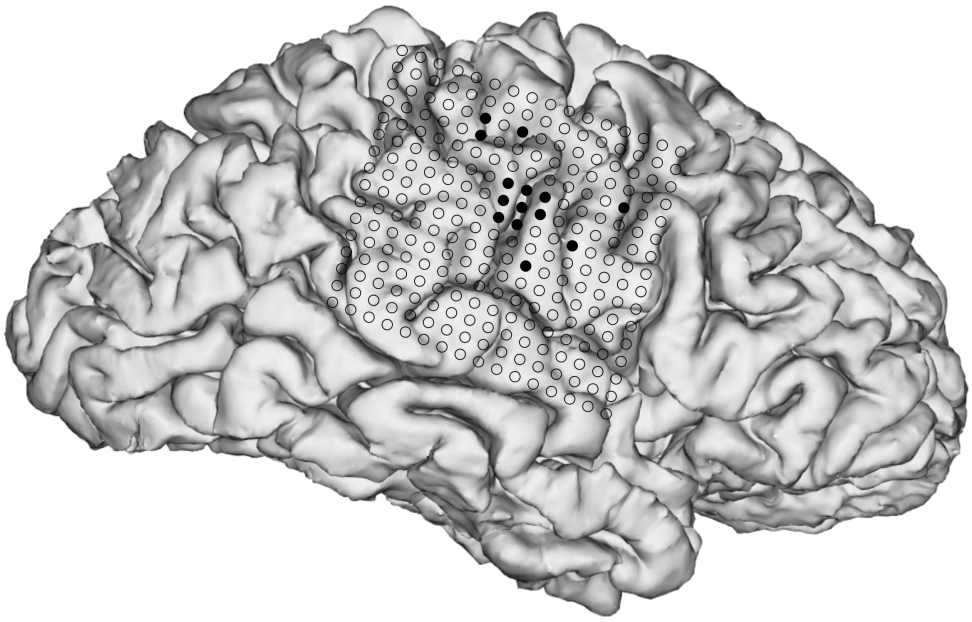

Figure 2.9: Electrode map of EC56 showing electrode locations that were selected by ANOVA ($p<0.05$)

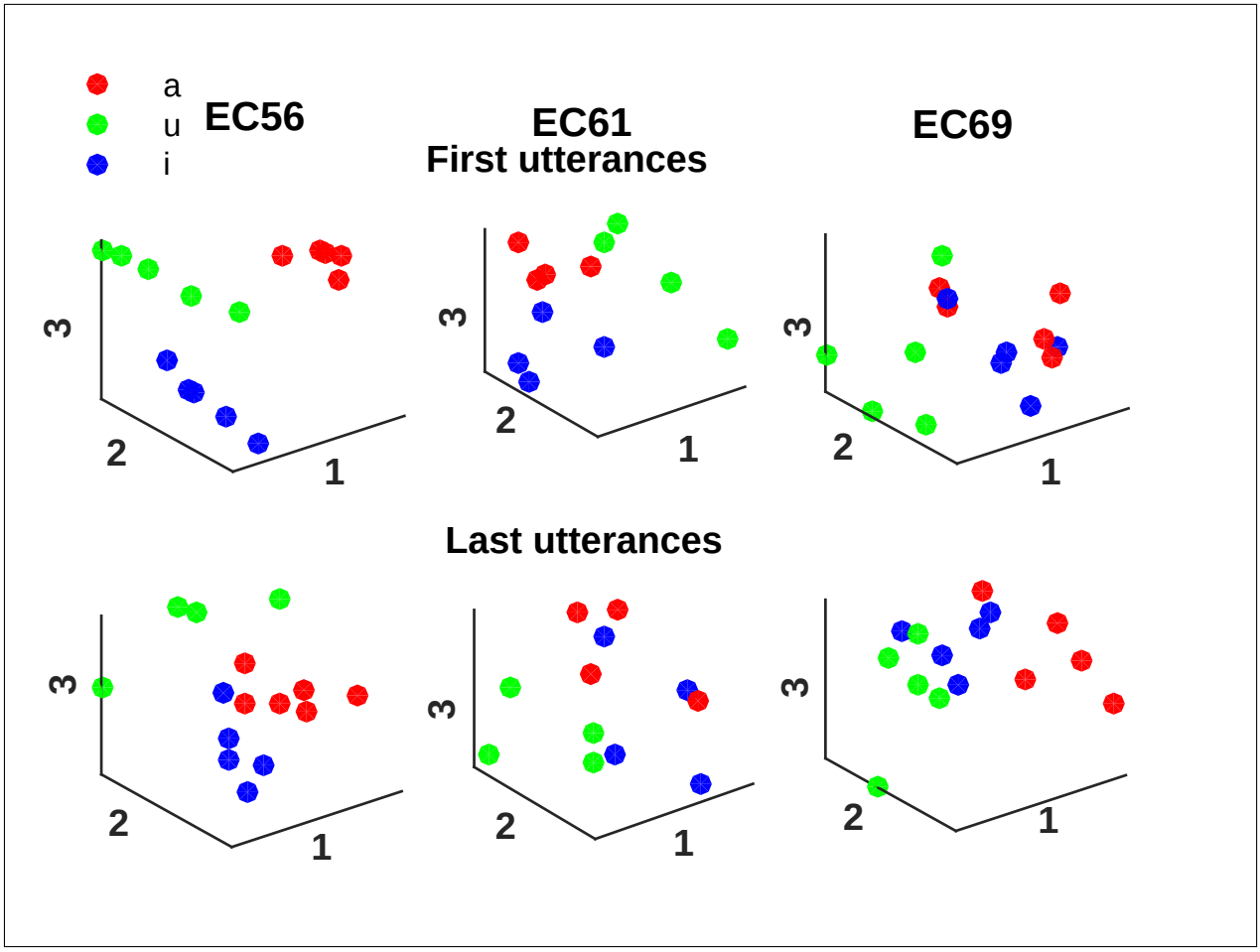

Figure 2.10: First three dimensions of features projected onto LDA subspace

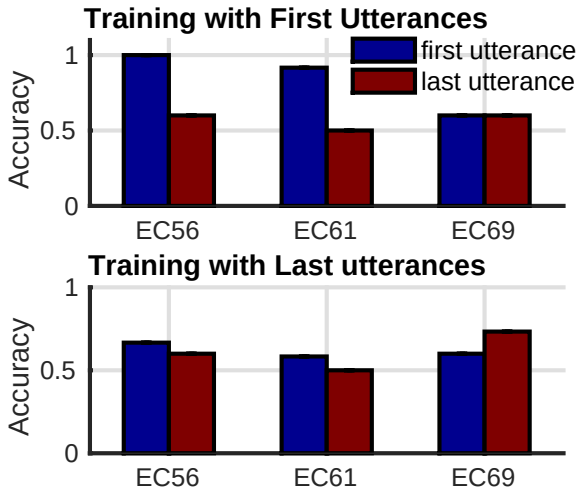

Figure 2.11: Comparing leave-one-out cross validation results from models trained with first utterances and last utterances across three subjects

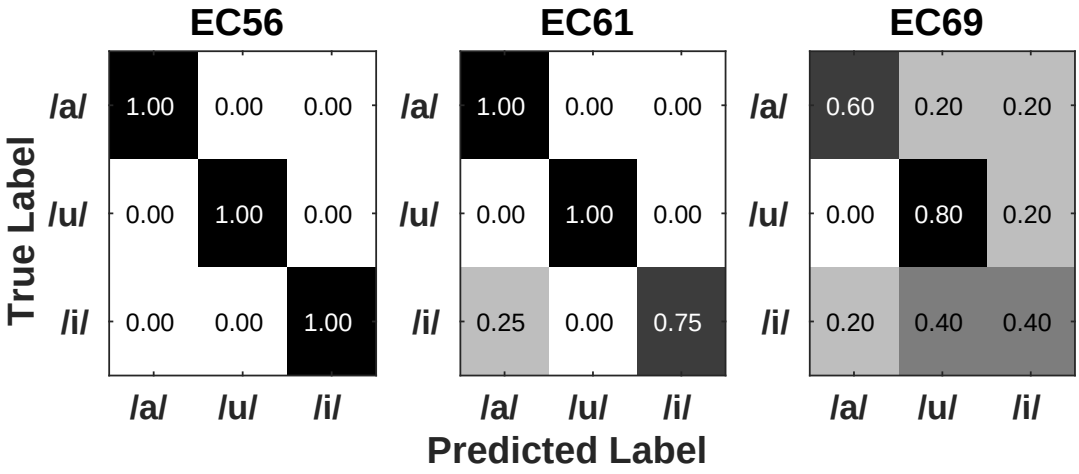

Figure 2.12: Confusion matrices from first utterance classification using model trained with trials from first utterance

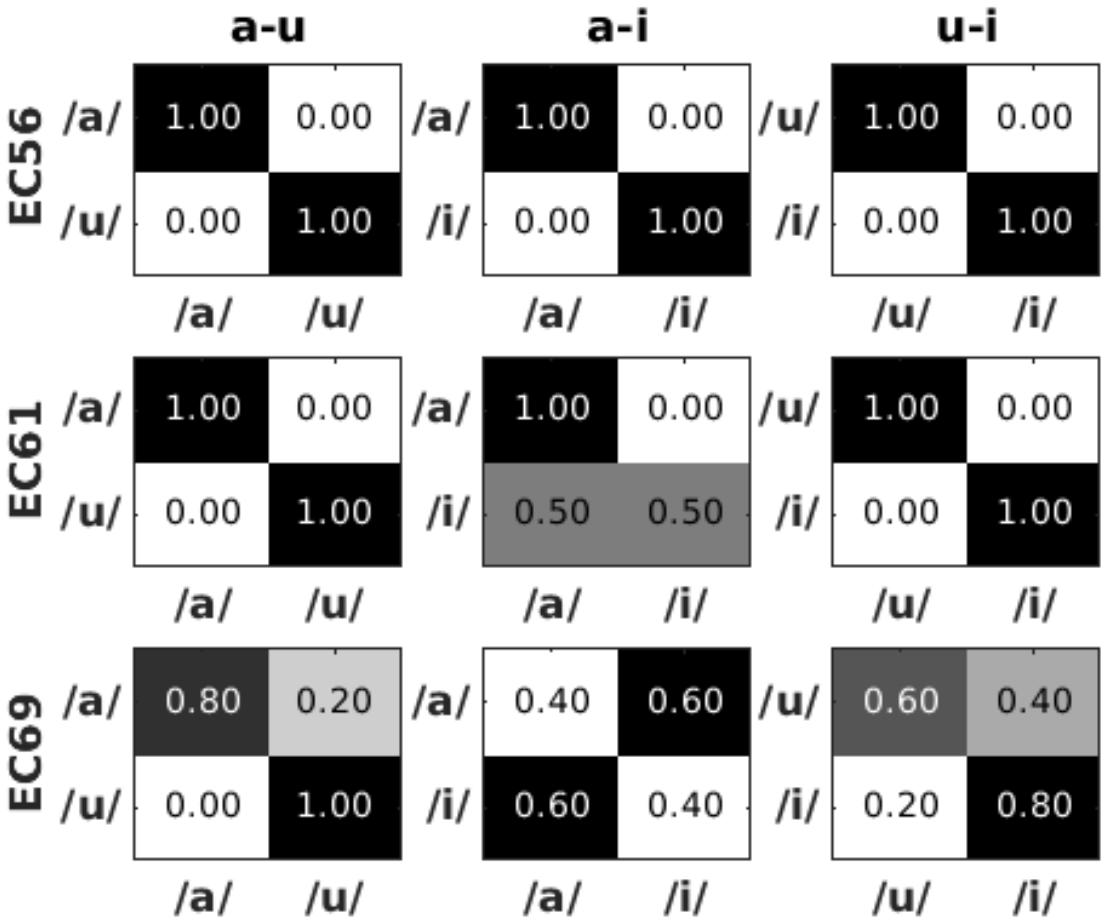

Figure 2.13: Confusion matrices for pairwise classification using trials from first utterance

three phonemes are used (See TABLE 2.3). In cases like this, it may be beneficial to use only two target phonemes to potentially achieve higher data rates.

Chapter 2, in part, is a reprint of the material from material as it appears in the Proceedings of the 2016 Engineering in Medicine and Biology Conference, "ECoG data analyses to inform closed-loop BCI experiments for speech-based prosthetic applications." Pailla, Tejaswy, Werner Jiang, Benjamin Dichter, Edward F. Chang, and Vikash Gilja. The dissertation author was the primary investigator and author of this paper. Chapter 2 also contains analyses done together with Anthony Au and is unpublished.

Table 2.3: Comparing pairwise phoneme classification with three phoneme classification

	Accuracy				ITC (in bpt)			
Subject	a-u	a-i	u-i	a-u-i	a-u	a-i	u-i	a-u-i
EC56	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.5843
EC61	1.00	0.75	1.00	0.9167	1.00	0.3270	1.00	1.3052
EC69	0.90	0.40	0.70	0.60	0.6180	0.0290	0.1246	0.4463

References

[1] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[2] Leigh R Hochberg, Mijail D Serruya, Gerhard M Friehs, Jon A Mukand, Maryam Saleh, Abraham H Caplan, Almut Branner, David Chen, Richard D Penn, and John P Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099):164, 2006.

[3] Leigh R Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y Masse, John D Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S Cash, Patrick Van Der Smagt, et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372, 2012.

[4] Eric C Leuthardt, Gerwin Schalk, Jonathan R Wolpaw, Jeffrey G Ojemann, and Daniel W Moran. A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, 1(2):63, 2004.

[5] J Adam Wilson, Elizabeth Felton, P Charles Garell, Gerwin Schalk, Justin C Williams, et al. Ecog factors underlying multimodal control of a brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):246–250, 2006.

[6] Ramesh Srinivasan, Paul L Nunez, and Richard B Silberstein. Spatial filtering and neocortical dynamics: estimates of eeg coherence. *Biomedical Engineering, IEEE Transactions on*, 45(7):814–826, 1998.

[7] Gerwin Schalk, KJ Miller, NR Anderson, JA Wilson, MD Smyth, JG Ojemann, DW Moran, JR Wolpaw, and EC Leuthardt. Two-dimensional movement control using electrocorticographic signals in humans. *Journal of neural engineering*, 5(1):75, 2008.

[8] Thilo Hinterberger, Guido Widman, Thomas Navin Lal, Jeremy Hill, Michael Tangermann, Wolfgang Rosenstiel, Bernhard Schölkopf, Christian Elger, and Niels Birbaumer. Voluntary brain regulation and communication with electrocorticogram signals. *Epilepsy & Behavior*, 13(2):300–306, 2008.

[9] Eric C Leuthardt, Charles Gaona, Mohit Sharma, Nicholas Szrama, Jarod Roland, Zac Freudenberg, Jamie Solis, Jonathan Breshears, and Gerwin Schalk. Using the electrocorticographic speech network to control a brain–computer interface in humans. *Journal of neural engineering*, 8(3):036004, 2011.

[10] Xiaomei Pei, Eric C Leuthardt, Charles M Gaona, Peter Brunner, Jonathan R Wolpaw, and Gerwin Schalk. Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *Neuroimage*, 54(4):2960–2972, 2011.

[11] Spencer Kellis, Kai Miller, Kyle Thomson, Richard Brown, Paul House, and Bradley Greger. Decoding spoken words using local field potentials recorded from the cortical surface. *Journal of neural engineering*, 7(5):056007, 2010.

[12] Xiaomei Pei, Dennis L Barbour, Eric C Leuthardt, and Gerwin Schalk. Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans. *Journal of neural engineering*, 8(4):046028, 2011.

[13] Kristofer E Bouchard and Edward F Chang. Neural decoding of spoken vowels from human sensory-motor cortex with high-density electrocorticography. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 6782–6785. IEEE, 2014.

[14] Emily M Mugler, James L Patton, Robert D Flint, Zachary A Wright, Stephan U Schuele, Joshua Rosenow, Jerry J Shih, Dean J Krusienski, and Marc W Slutzky. Direct classification of all american english phonemes using signals from functional speech motor cortex. *Journal of neural engineering*, 11(3):035015, 2014.

[15] Timothy Blakely, Kai J Miller, Rajesh PN Rao, Mark D Holmes, and Jeffrey G Ojemann. Localization and classification of phonemes using high spatial resolution electrocorticography (ecog) grids. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4964–4967. IEEE, 2008.

[16] Kristofer E Bouchard, Nima Mesgarani, Keith Johnson, and Edward F Chang. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441):327–332, 2013.

[17] Eric C Leuthardt, Kai J Miller, Gerwin Schalk, Rajesh PN Rao, and Jeffrey G Ojemann. Electrocorticography-based brain computer interface-the seattle experience. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):194–198, 2006.

[18] Stephen I Ryu and Krishna V Shenoy. Human cortical prostheses: lost in translation? *Neurosurgical focus*, 27(1):E5, 2009.

[19] Nicholas G Hatsopoulos and John P Donoghue. The science of neural interface systems. *Annual review of neuroscience*, 32:249, 2009.

[20] Kai J Miller, Eric C Leuthardt, Gerwin Schalk, Rajesh PN Rao, Nicholas R Anderson, Daniel W Moran, John W Miller, and Jeffrey G Ojemann. Spectral changes in cortical surface potentials during motor movement. *The Journal of neuroscience*, 27(9):2424–2432, 2007.

[21] P. Shenoy, K.J. Miller, J.G. Ojemann, and R.P.N. Rao. Generalized features for electrocorticographic bcis. *Biomedical Engineering, IEEE Transactions on*, 55(1):273–280, Jan 2008.

[22] Kai J Miller, Pradeep Shenoy, John W Miller, Rajesh PN Rao, Jeffrey G Ojemann, et al. Real-time functional brain mapping using electrocorticography. *Neuroimage*, 37(2):504–507, 2007.

[23] Pradeep Shenoy, Kai J Miller, Jeffrey G Ojemann, and Rajesh PN Rao. Finger movement classification for an electrocorticographic bci. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pages 192–195. IEEE, 2007.

[24] Justin C Sanchez, Aysegul Gunduz, Paul R Carney, and Jose C Principe. Extraction and localization of mesoscopic motor control signals for human ecog neuroprosthetics. *Journal of neuroscience methods*, 167(1):63–81, 2008.

[25] Walter J Freeman. Hilbert transform for brain waves. *Scholarpedia*, 2(1):1338, 2007.

[26] Nuri F Ince, Fikri Goksu, and Ahmed H Tewfik. An ecog based brain computer interface with spatially adapted time-frequency patterns. In *BIOSIGNALS (1)*, pages 132–139, 2008.

[27] Onder Aydemir and Temel Kayikcioglu. Wavelet transform based classification of invasive brain computer interface data. *Radioengineering*, 20(1):31–38, 2011.

[28] G Charvet, M Foerster, S Filipe, J Porcherot, JF Bêche, R Guillemaud, P Audebert, G Régis, B Zongo, S Robinet, et al. Wimagine: A wireless, low power, 64-channel ecog recording platform for implantable bci applications. In *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, pages 356–359. IEEE, 2011.

[29] Rizwan Bashirullah. Wireless implants. *Microwave Magazine, IEEE*, 11(7):S14–S23, 2010.

[30] Bernhard Graimann, Jane E Huggins, Simon P Levine, and Gert Pfurtscheller. Toward a direct brain interface based on human subdural recordings and wavelet-packet analysis. *Biomedical Engineering, IEEE Transactions on*, 51(6):954–962, 2004.

[31] Chih-Wei Hsu and Chih-Jen Lin. A comparison of methods for multiclass support vector machines. *Neural Networks, IEEE Transactions on*, 13(2):415–425, 2002.

[32] Kai-Bo Duan and S Sathiya Keerthi. Which is the best multiclass svm method? an empirical study. In *Multiple Classifier Systems*, pages 278–285. Springer, 2005.

[33] Christopher M Bishop et al. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[34] John J Ohala. The physiology of tone. *Southern California occasional papers in linguistics*, 1:1–14, 1973.

[35] Andreas Bruns. Fourier-, hilbert-and wavelet-based signal analysis: are they really different approaches? *Journal of neuroscience methods*, 137(2):321–332, 2004.

[36] Partha P Mitra and Bijan Pesaran. Analysis of dynamic brain imaging data. *Biophysical journal*, 76(2):691–708, 1999.

[37] T Mullen. Cleanline eeglab plugin. *San Diego, CA: Neuroimaging Informatics Toolsand Resources Clearinghouse (NITRC)*, 2012.

[38] David Conant, Kristofer E Bouchard, and Edward F Chang. Speech map in the human ventral sensory-motor cortex. *Current opinion in neurobiology*, 24:63–67, 2014.

[39] Pascal O Vontobel. A generalized blahut-arimoto algorithm. In *Information Theory, 2003. Proceedings. IEEE International Symposium on*, page 53. IEEE, 2003.

Chapter 3

Autoencoders for learning template spectrograms in ECoG signals

3.1 Introduction

Electrocorticography (ECoG) is a neural recording technique wherein electrical potentials are measured from the surface of cortex. One of the primary clinical applications of ECoG is to evaluate epileptic activity: individuals with chronic intractable epilepsy are implanted with ECoG electrodes and their cortical activity is monitored in an epilepsy monitoring unit (EMU) to localize epileptic seizure foci. In this setting, the implanted electrodes are also used to provide electrical stimulation to map functional brain areas. These maps guide surgical resection of affected brain areas while attempting to minimize the potential for functional deficits. Patients in the EMU are often recruited to take part in neurophysiology studies and ECoG data are collected while they perform cognitive and behavioral tasks. Given this unique opportunity to study the human brain in action, such studies have significantly contributed to the advancement of human neuroscience research, particularly with respect to two domains: functional brain mapping [1, 2] and brain-computer interface (BCI) development [3].

Most functional brain mapping studies analyze spatio-temporal patterns of ECoG signal modula-

tion in the 8-40 Hz band or the 70-200 Hz band or sub-bands in those ranges to identify cortical processes underlying behaviors [2, 4–7]. Precise labeling of behavioral onset and offset timings is important when performing such analyses since event related potentials are sensitive to the timing of activity [8]. However, many of the labeling methods employed in experimental studies rely upon the subjects' reaction times to a cue and this can potentially introduce errors in analysis; thus, while ECoG may afford high temporal resolution, study design and conventional analysis methods can limit the temporal resolution of interpretations.

The high spatio-temporal resolution and signal-to-noise ratio (SNR) of ECoG also makes it a promising source for BCI control signals. Several studies have explored the use of ECoG for prosthetic control [3, 9–12] and ECoG based BCI systems are being explored for long term use as communication prosthesis for individuals with locked-in syndrome [13]. Evidence of ECoG signal modulation in specific spectral bands, such as 8-32 Hz and 76-100 Hz [14] and in broad spectral changes in the 5-200 Hz frequency range [15] during motor activity, has motivated the majority of ECoG based BCI studies to use frequency specific spectral power as features for decoding. These spectral powers are typically estimated from a specific time window around activity onset. However, the typically applied conventional and rigid feature selection of frequency bands and time windows may limit BCI system performance, since the precise range of informative frequency bands and temporal range of relevant modulation may vary across subjects [16]. Also, the activation patterns in different cortical areas for the same task might vary for the same subject throughout the duration of the experiment as the subject's familiarity with the task changes [17]. Furthermore, computing spectral powers in aggregated time bins does not fully exploit the temporal resolution that ECoG provides. With conventional approaches, decoding ECoG signals with high accuracy requires a neural engineer to carefully analyze spatial, temporal and frequency patterns in the data and hand design custom subject-specific, task-specific features and decoding models, resulting in a herculean task.

Thus, the complexity of feature engineering motivates the need for automated pattern recognition tools that can efficiently extract spatial, temporal and frequency patterns in ECoG data. In recent years,

several neuroimaging and EEG based studies have used Deep Learning (DL) methods for pattern recognition and classification tasks [18–23]. Deep belief networks (DBN) have been applied to learn neural correlates in real-time for an ECoG based BCI [24]. In another study [25], a variant of convolutional stacked autoencoder [26] has been used to extract hierarchical features from ECoG signals. Other studies [27–30] have used echo state network implementations of recurrent neural networks (RNNs) to model neural activity for use in BCIs. These studies demonstrate that deep neural network architectures can achieve performance comparable to the state-of-the-art achieved by conventional machine learning methods and, in addition, have the potential to extract physiologically relevant features [18]. However, most studies that use deep learning methods to decode neural data use network architectures that are successful in other domains, mainly computer vision (For example, [31] used a network mimicking the VGG architecture that was developed for object recognition). Although such architectures achieve good performance results, the number of parameters to be learnt is on the order of millions, thus risking overfitting. Given the small dataset sizes in neurophysiology studies, network architectures designed with domain knowledge could address the problem of overfitting and improve generalizability of architectures across subjects and tasks.

In this work, we learn generalized time-frequency (t-f) patterns, which we call "template spectrograms" (TS), that can effectively encode patterns from ECoG electrode channels at different spatial locations using autoencoders (AE) [32]. These patterns find representations that are robust to temporal and frequency shifts in the data. We show the efficacy of TS in two contexts: BCI decoding and identification of functionally similar cortical regions (Figure 3.1). We use a convolutional neural network framework that uses TS to decode individual finger flexions, i.e. identify which finger is moving, from ECoG data and demonstrate classification performance that surpasses models that achieved when using traditional spectral features. We also evaluate the generalizability of TS by using a subject-to-subject transfer learning approach. The results suggest that for this finger flexion task, informative neural patterns across brain regions and subjects can be summarized using a limited number of TS. To investigate potential functional mapping applications, we

use learnt t-f representations to cluster functionally similar brain regions and validate clustering results by comparing them with prior knowledge of electrode location. We observed that brain areas that have task relevant responses are clustered together. This suggests that unsupervised learning based methods could aid in the mapping of brain areas to behaviors.

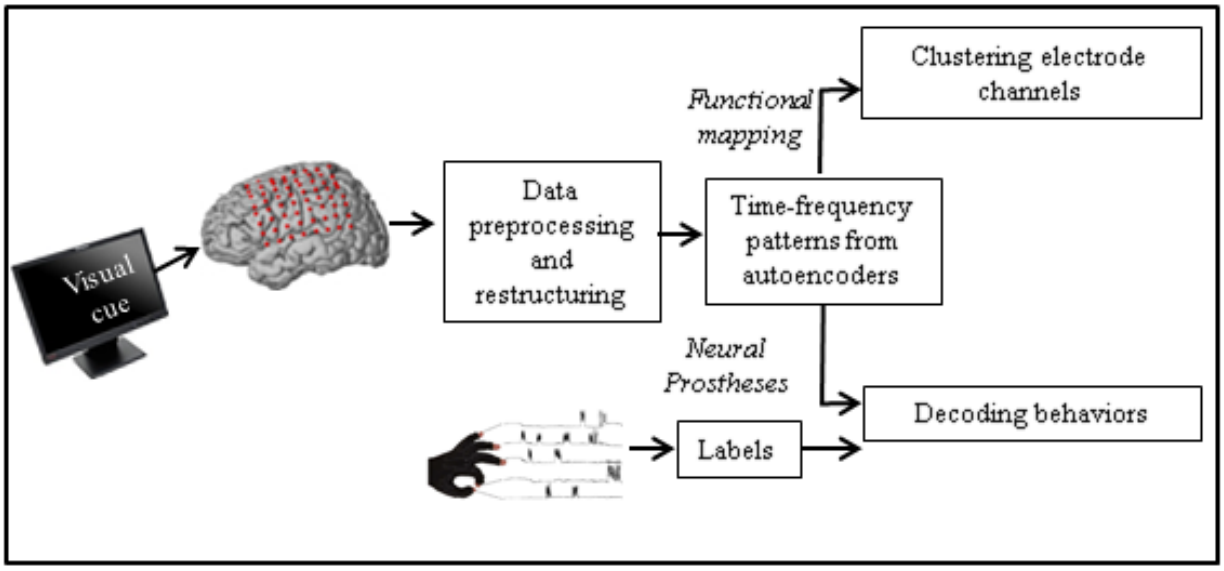

Figure 3.1: Schematic of proposed electrocorticographic (ECoG) data analysis approach. ECoG signals and behaviors are synchronously recorded while the subject performs a queued tasks. In the primary datasets considered for this study, a finger flexion task was utilized.

3.2 Methods

3.2.1 Recordings

Data were collected from 9 subjects with intractable epilepsy who underwent temporary placement of subdural electrode arrays to localize seizure foci prior to surgical resection. All the subjects participated in a purely voluntary manner, after providing informed written consent, under experimental protocols approved by the Institutional Review Board of the University of Washington (#12193). All patient data were anonymized according to IRB protocol, in accordance with HIPAA mandate. These data originally appeared in [33] and were publicly released in [34]. Three subjects' data were excluded from the main analyses due to

inconsistencies in recorded behavior and low trial count (see section 3.2.2 for more details about exclusion criteria). Subjects were instructed to move specific individual fingers of the hand contralateral to the implant in response to visual cues. During the finger movement task, subjects were cued with a word displayed on a bedside monitor indicating which finger to move during 2 second movement trials. They performed self-paced movements in response to each of these cues, and typically moved each finger 2-5 times during each trial, but some trials included many more movements. A 2 second rest trial (blank screen) followed each movement trial. There were 30 movement cues for each finger, and trials were interleaved randomly. Finger positions were recorded using a data glove (5DT Data Glove 5 Ultra, Fifth Dimension Technologies; see Figure 3.2 for example data glove recordings for subject I). Electrode locations were determined using localization methods described in [35] and were also manually verified.

3.2.2 Subject exclusion criteria

The task required subjects to flex fingers for two seconds and rest for two seconds, thus allowing for neural activity to return to baseline during the intertrial interval. However, not all subjects followed the cues. Figure 3.3 below shows a snapshot of normalized, detrended flexion traces for subjects 'A-F' and 'G-I' respectively for a period of 30 seconds. Flexions were inconsistent across subjects and trials. For example, subject 'C' presented in the main results did not rest during the two second intertrial period.

The decoding results for subjects 'G-I' are excluded in the main results, but are shown in Figure 3.14. For each subject, the {number of SMC electrodes}/{total number of electrodes} is 8/58, 12/43, 3/38 for subjects G-I respectively. These subjects were excluded from the main manuscript due to inconsistent flexion behavior and low trial count (¡250).

- Subject G's flexion behavior was inconsistent (similar to subject C Fig 3.4.) and inspection of glove data shows overlapping flexion traces from multiple fingers. Trials labeled using the method described in section above couldn't be verified with expected behavior. (Unlike subject C's trial labels which

could be manually verified, which allowed us to include subject C). This discrepancy could either stem from faulty data glove or due to subject not following cues. We excluded this subject owing to this ambiguity in labelling.

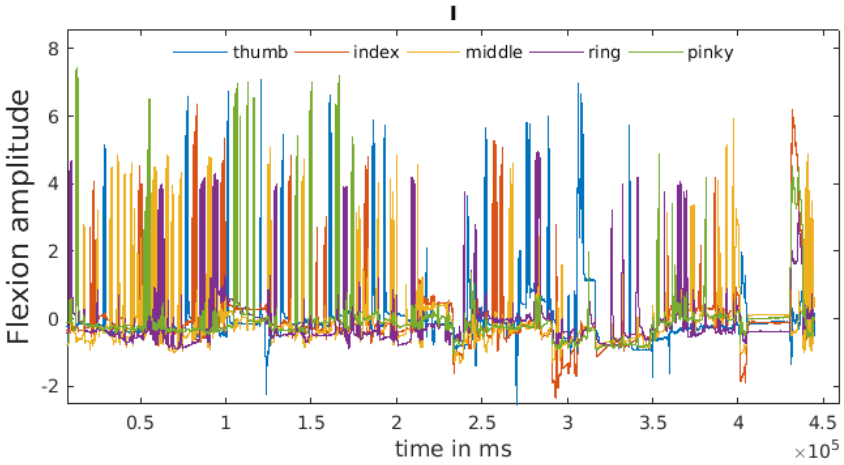

Figure 3.2: Complete flexion trace for subject 'I'.

- Subject H's experiment duration was less than 3 minutes resulting in very few trials per finger and Subject I had highly variable flexion duration across trials and appeared to not follow the task mid experiment and not flex their fingers for a long enough duration to have a sufficient number of 600 ms trials for training AE (See Figure 3.2). There were less than 160 total trials for subjects H and I.

3.2.3 Preprocessing

The ECoG signals were amplified, bandpass filtered between 0.15 and 200 Hz, digitized at 1000 Hz and stored along with the time synchronized digitized flexion samples for all fingers. Noise common to all channels was removed by common average referencing. Signals from each channel were band pass filtered with windowed sinc type I linear phase filters of 4 Hz bandwidth between 4-140 Hz (i.e. 4-8, 8-12 .... 136-140 Hz). After excluding bands that include line noise and its harmonic frequencies (i.e. 56-60, 60-64, 64-68, 116-120, 120-124, 124-128), 28 frequency bands were used. Filtered signals were z-scored.

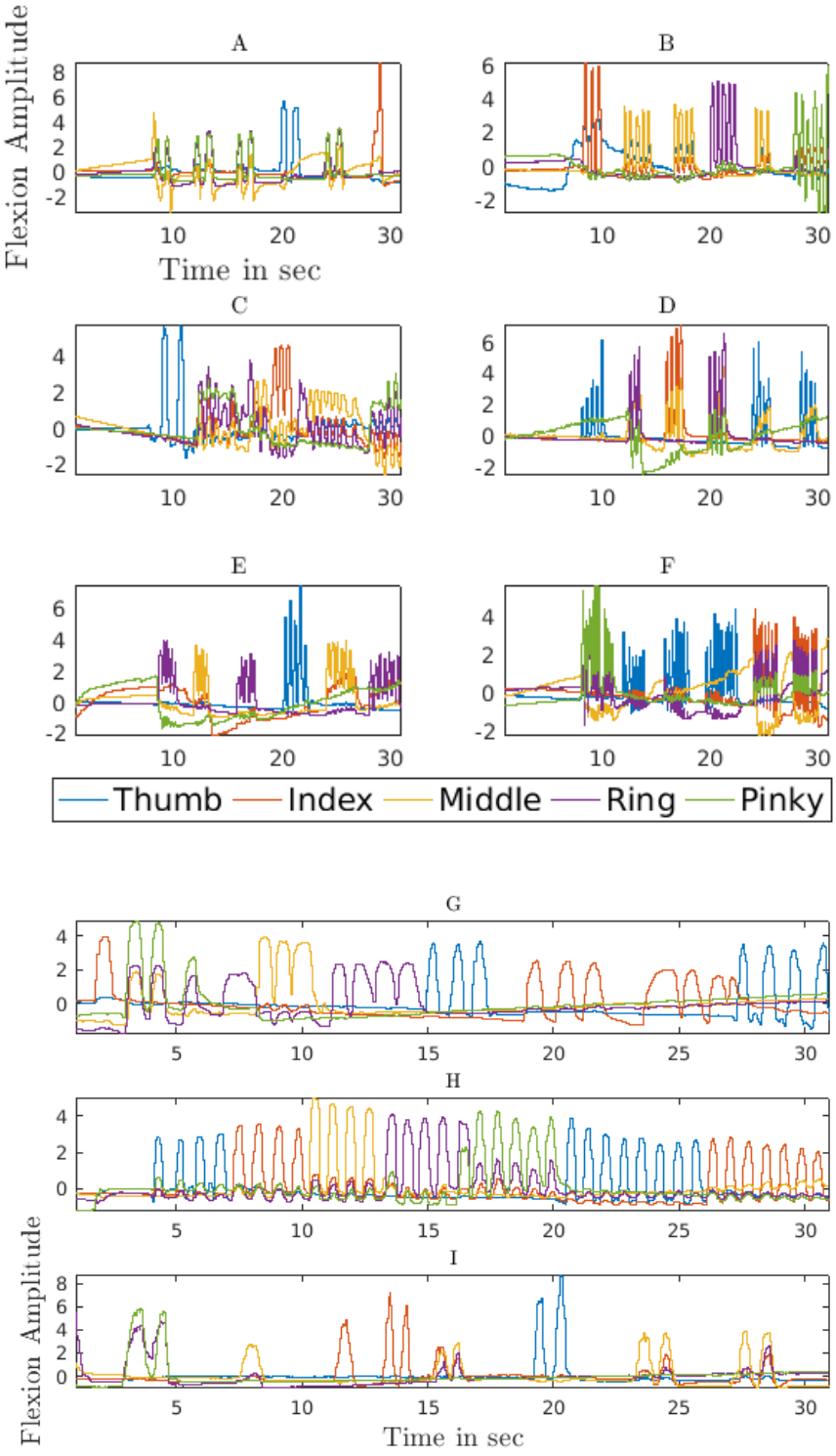

Figure 3.3: Flexion traces of sample 30 sec of subjects A-F presented in main results (top) and excluded subjects G-I (below).

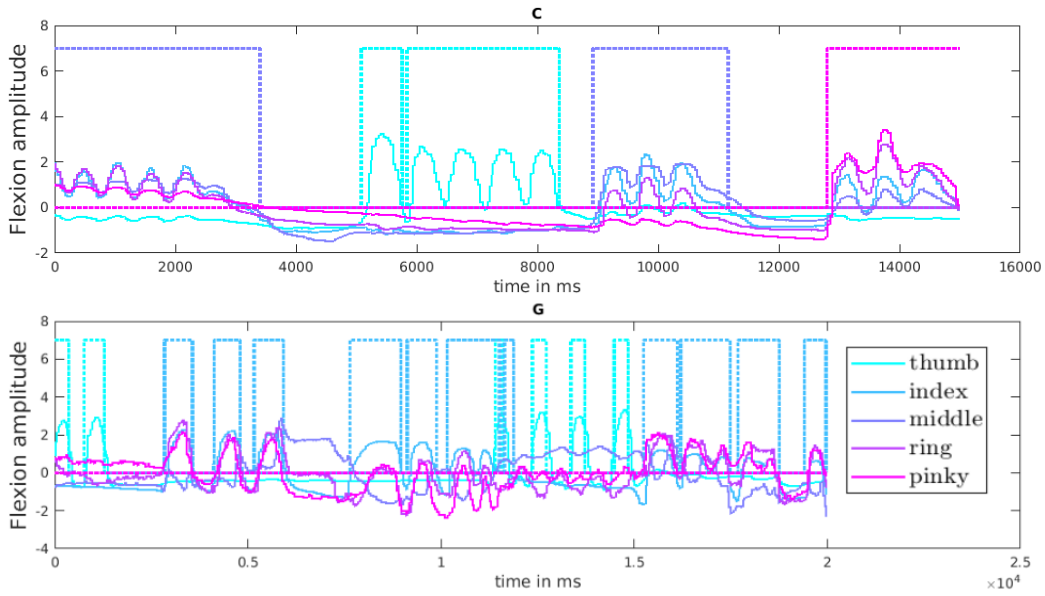

Figure 3.4: Sample flexion trace for subject C ad G

The flexion periods were further broken into 600 ms non overlapping windows which we define as trials for our analysis. Rest data from inter trial intervals were not used. The 600 ms blocks were split into 6 bins of 100 ms each and signal power in each frequency band and each 100 ms time bin was estimated. For all our analysis, to estimate the power of a signal for a frequency band and time bin, we calculated the average of the squared amplitude of the bandpass filtered signal within the time bin. Accordingly, from each electrode channel, trial blocks were restructured into a 28×6 matrix (see Figure 3.6a) in which each entry is the estimated power for a specific frequency band and time bin. Neural signals were delayed by 150 ms to account for delay between cortical activity and motor movement. Each matrix was labeled with the finger that was flexed during the corresponding block of time, as determined by analyzing the glove data. The method we used for labeling trials is described in section 3.2.4 below. The number of 600 ms trials for each subject are 278, 464, 364, 291, 541 and 463, for subjects A-F, respectively.

3.2.4 Labelling Trials

Flexion traces from data gloves are first detrended to eliminate baseline drift in sensor readings and zscored to normalize signal amplitude. Thresholds for detecting finger flexion onset are set as follows. Let x be the time series comprising the flexion trace of a finger. The values in x are sorted into 10 equally spaced bins between the minimum and maximum values of x. Let p be the bin center with the maximum count. The threshold for detecting finger movement is set as $\frac{2}{5}(max(x) - p) + p$. Movement onset is detected when the flexion trace crosses the threshold (Fig 3.5). Consecutive detected movements are merged into a single trial if they occur within a set time difference between each other (determined empirically). If two finger movements are detected during the same period, the finger cued by the stimulus is marked as flexed.

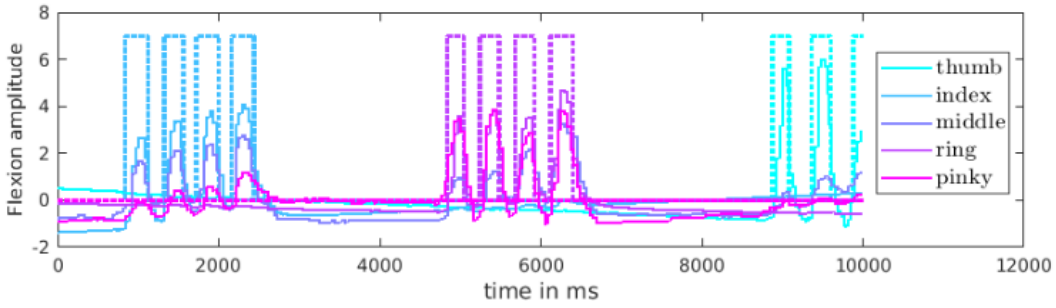

Figure 3.5: Sample flexion trace from Subject D with dotted lines showing labelled trials.

3.2.5 Autoencoders for learning time-frequency patterns

An autoencoder, in its simplest form, is a feed-forward neural network that takes an input and maps it to a hidden layer (encoder). The latent representation, i.e. the output from hidden layer is then mapped back (decoder) into a reconstruction of the same shape as input. The weights of encoder and decoder layers are optimized to minimize error between input and reconstructed output [32, 36]. Autoencoders thus learn efficient representation of data and are commonly used for unsupervised feature extraction and dimensionality reduction in neural networks [37, 38].

We use AEs to find t-f patterns (or kernels) in ECoG data. Restructured data from each electrode channel is given as a separate input to an AE and h t-f kernels are learnt using a saturated linear ($satlin$) activation function in Eq. 3.1, where h is the number of hidden nodes in AE. We expect the $satlin$ activation to make the model robust to artifacts by saturating extreme activations that might result from noisy data.

$$f(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ z & \text{if } 0 < z < 1 \\ 1 & \text{if } z \geq 1 \end{cases} \tag{3.1}$$

The rationale behind this AE approach is that each channel might have its own task relevant t-f pattern and we want to find a set of templates which have representations generalizable for all channels. To determine the sparsity proportion, i.e. the average output of each hidden unit over the training set, we examined the mean squared error between reconstructed and input signals for varying levels of sparsity proportion and h, and observed that there was no significant trend. So, we used a sparsity proportion of 0.05 for all our analyses. AEs are trained to minimize the sparse mean squared error with l_2 weight regularization for a maximum of 200 epochs. The encoder of a trained AE can be used to generate new t-f based features for neural decoding and in a neural network based decoding architecture (see Section 3.2.6 and Figure 3.6b) the encoder can be directly integrated into the architecture. Figure 3.7a shows example t-f patterns learnt for a subject. We evaluate the efficacy of these representations in two applications: BCI decoding and functional mapping.

3.2.6 Deep neural networks for decoding

Deep neural networks (DNNs) learn hierarchical representations of data [39] and provide a way to embed feature abstraction as a part of goal-directed learning. In this work, we use a 3-layer AE initialized DNNs (AE-DNNs) to decode individual finger flexions. The architecture is shown in Figure 3.6b and is

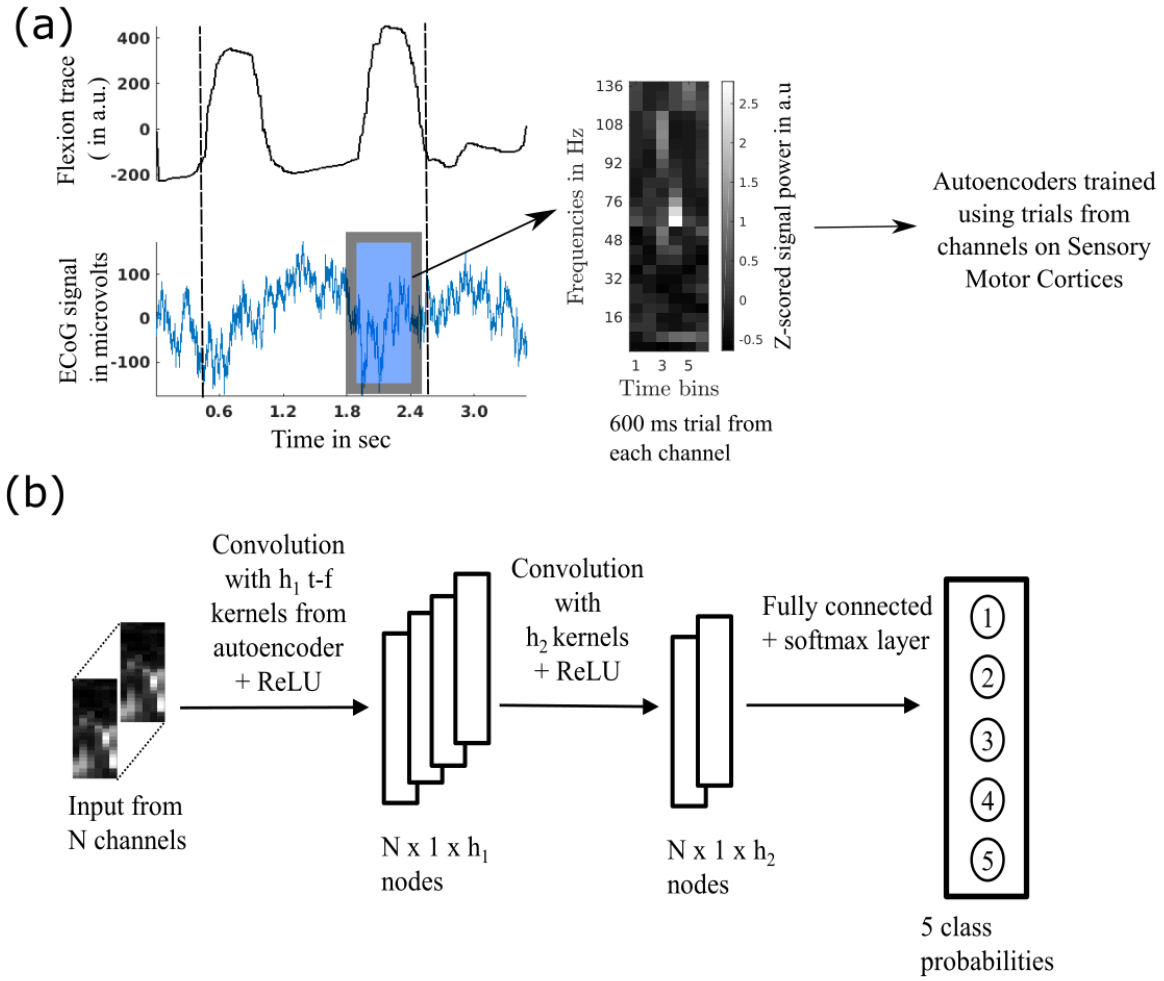

Figure 3.6: Schematic of network architecture. (a): Trials from each channel are restructured in time-frequency (t-f) format and autoencoders are trained to extract t-f patterns. ECoG signal (in blue, shown for a single channel) is synchronized with finger flexion trace (in black). The dashed lines represent the start and end of a flexion period.(b): t-f kernels learnt by autoencoders are used as weights in DNNs to decode finger flexions. DNN takes input from N channels. Weights for each electrode channel, i.e. spatial modeling, are learnt in the last layer.

implemented with a convolutional framework. N denotes the number of electrode channels for each subject and h_i denotes the number of hidden nodes in each layer. The first layer's weights and biases are fixed with TS learnt from an AE that transforms data from each channel to h_1 nodes. The second layer transforms each of the h_1 nodes in the first layer to h_2 nodes and the last layer is a fully connected softmax layer which maps $N \times h_2$ nodes to 5 nodes that give class probabilities. So, we are effectively learning the discriminative

spatial pattern in the last layer. We use a rectified linear unit (ReLU) non-linearity between the layers as we observed that the training was more stable as compared to training with *sigmoid* and *tanh* non-linearities. Unless otherwise specified, we only train the last two layers of DNN while keeping the first layer weights from AE fixed. We perform 5-fold cross validation (CV) for evaluating decoding performance. The trials are first randomized and DNNs are trained on 4/5th of the data and tested on the remaining 1/5th. While randomizing trials, we ensured that possibly correlated 600 ms trials split from the same flexion epoch are either in training set or in test set. We also ensured that the TS are obtained from AE using only the trials from training set. The network with the minimum validation loss is chosen to decode trials in the test set. Sample scripts used for decoding analyses are available at `https://github.com/tejaswy/ECoG-TF`. Neural network training is done with MatConvNet [40]. For decoding results, we used electrodes from sensory motor cortices (SMC) which includes dorsal M1, dorsal S1 and ventral SMC regions. For each subject, the {number of SMC electrodes}/{total number of electrodes} is 11/46, 11/64, 15/63, 12/47, 11/61 and 9/64 for subjects A-F respectively. Selection of hyperparameters, such as the number of nodes in each neural network layers, is made using a grid search. A grid search between 10 to 100 with a step size of 10 was used to select the number of hidden nodes in AE. The same grid search was completed for choosing the number of principal components. The number of nodes in the second layer of the DNN was selected from a grid search between 5 to 40 with a step size of 10. The models which minimized cross entropy loss was selected. For training neural networks, the learning rate for stochastic gradient descent algorithm was set to a linearly decreasing value from 0.025 to 0.001 with a step size of 0.001. The maximum number of training epochs was fixed to 25 and batch normalization was adopted with a batch size of 10. Training is done with a training to validation split of 4:1.

Transfer learning

Limited ECoG dataset size for structured tasks is a major impediment to applying deep learning methods for BCI studies. We show that using a transfer learning (TL) approach, we can leverage the representations learnt across subjects to decode flexions for a new subject. We use data from 4 subjects to learn TS and use them in the first layer of DNN to decode flexions of fifth subject. The rationale behind this is that there may be generalizable t-f patterns in different brain areas across subjects and extracting these patterns from the available data corpus and using them for a new subject by learning the spatial structure by training the last layers might be advantageous in data limited conditions. Thus, the goal for the AE is to effectively learn a "library" of relevant TS.

3.2.7 Baseline models

Spectral features (SF), i.e. estimated spectral powers in low (8-32 Hz) and high (76-100 Hz) frequency bands are prominently used as features in ECoG based BCIs [14, 41]. Linear Discriminant Analysis (LDA) classifier is commonly used in ECoG based BCI studies [11]. Thus, to evaluate the efficacy of different features, i.e. SF, time-frequency powers described in Section 3.2.3 (TF), and AE, as well as to evaluate different decoding models, i.e. LDA vs. DNN, we compare decoding results for the following pairings of features and decoding models.

- **SF-LDA**: We use SF calculated in 6 bins of 100 ms each with a regularized LDA classifier. To calculate SF, we bandpass filter signals in 8-32 Hz and 76-100 Hz band and calculate the spectral power estimate using the procedure described in Section 3.2.3. So, for a given channel, we have 12 SF per trial. This provides a baseline model with traditional SF and a linear decoding model.

- **TF-LDA**: We used estimated powers in time-frequency arrays and a regularized LDA classifier. This model will demonstrate if using information from the entire frequency range and smaller frequency

bins improves decoding performance when compared to SF.

- **AE-LDA**: We used the t-f features learnt by AE and a regularized LDA classifier. This provides a baseline to compare SF with AE features.

- **CK-DNN**: We created custom kernels (CK) that act as t-f filters shown in Figure 3.7b by using an entry of 1 in rows that cover frequency bands from 8-32 Hz and/or 76-100 Hz and 0 for the rest. We have combinations of kernels ranging all temporal scales i.e. 1 to 6 time bins. This give us a total of 63 kernels. We use these custom kernels in the first layer of the DNN described previously for classification. We use this baseline comparison model to be consistent with conventionally used SF while providing the advantage of a DNN for decoding.

- **PCA-LDA**: The weights of an autoencoder with linear activation and lower number of hidden units (n) than number of inputs can span the same subspace as the first n eigen vectors of the covariance matrix of features. Hence, Principal Component Analysis (PCA) could be a reasonable alternative to the autoencoder based feature extraction methods, specifically in data limited conditions.

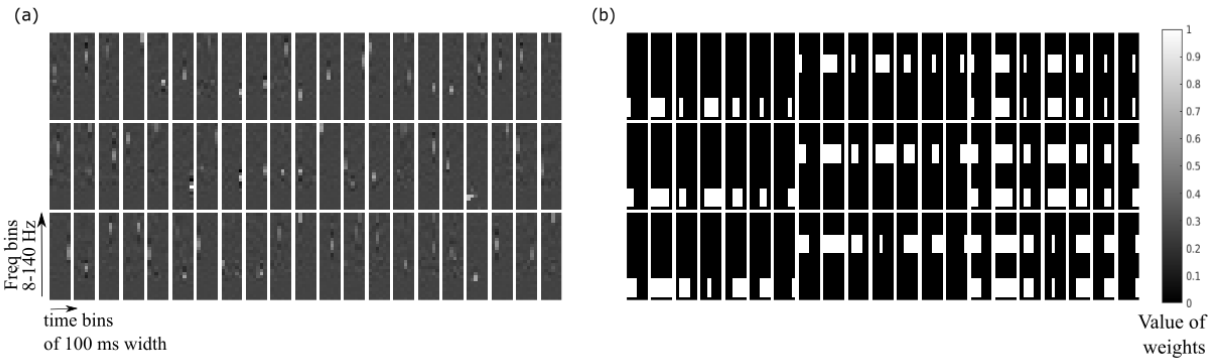

Figure 3.7: Comparison of t-f kernels. Kernel ordering in this figure is arbitrary. Axes are frequency vs time bins as shown in Figure 3.6a. (a) An example of 63 kernels learnt by autoencoders (b) 63 custom kernels used in CK-DNN.

3.2.8 Regularization for LDA

We used a regularized LDA where the class covariance is defined by $\Sigma_\gamma = (1-\gamma)\Sigma + \gamma diag(\Sigma)$, where Σ is the empirical, pooled covariance matrix and γ (where $\gamma \in [0,1]$) is the level of regularization. Optimal values of γ are obtained using Bayesian optimization.

3.2.9 Robustness of features in noisy conditions

Noise in ECoG recordings due to external sources is a common occurrence. Figure 3.8 shows a sample ECoG signal recorded in the epilepsy monitoring unit outside of controlled experimental conditions. Unlike in controlled experimental settings, brain-computer interfaces have to be robust to external noise when used in real world applications. System design choices, specially for invasive monitoring devices or implants, should also consider the trade-off between battery life and SNR. For example, using low-power amplifiers could result in longer battery life but comes at the expense of SNR. The use of signal processing pipelines robust to noise could play a crucial role in making such design choices.

Due to the use of a non-linear activation, we expect AE based features to be more robust to noise compared to PCA or TF features. To verify this, we simulated frequently occurring spike noise in ECoG by introducing broadband noise of varying levels in randomly selected trials in the training data. We then trained LDA and obtained predictions on test data.

3.2.10 Noise levels in PCA vs AE comparison

Noise due to external sources are commonly observed in ECoG recordings. To test the robustness of TF-LDA, PCA-LDA and AE-LDA to noise events commonly observed in ECoG recordings, we added simulated noise with varying amplitude to ECoG channels. Figure 3.8 shows signal from an ECoG channel recorded in the epilepsy monitoring unit. Common Average Referencing (CAR) is a commonly used technique to remove noise sources common to recorded channels. We performed CAR on this noisy channel

to get a cleaner signal devoid of spike noise. Power spectral densities (psd) shown gives us an estimate of the broadband power shift caused by the noise in clinical recording conditions. Drawing an estimate of scale of broadband power changes caused by noise, we added broadband noise to the time-frequency arrays (described in the Preprocessing section) of channels randomly selected from 30% of training samples. We

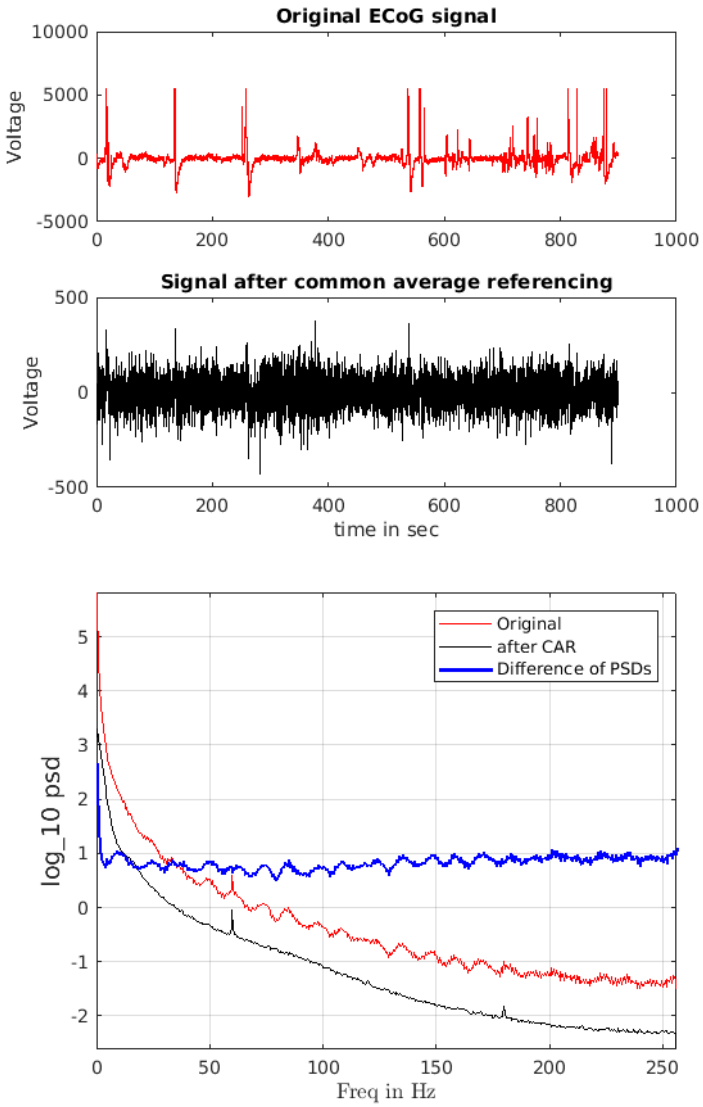

Figure 3.8: Effect of transient noise on ECoG signals. ECoG signal before common average referencing (top) shows the commonly occuring spike noise in ECoG recordings. Common average referencing removes majority of the spike noise (bottom left). The power spectral density plots (bottom) show broadband power shift occuring due to noise.

do this by generating a 1x28 array which is an estimate of noise psd following the Power-law $m.\frac{10^6}{f^2}.^1$. This noise addition is done prior to normalization of features across bands. Here the scaling factor m is varied from 1 to 5 to simulate increasing noise levels and f is the center frequency in each of 4 Hz frequency bins.

3.2.11 Generalization of autoencoder features

We tested our methods on a new dataset to see how well AE features generalize. We used a motor movement and imagery dataset where 5 subjects perform both actual movements of hand and tongue and kinesthetic imagery. During imagery task, subjects were asked to imagine making the movements. However there is no definitive way to verify if they are following the cues during imagery tasks. This dataset was first published in [42]. We used all recorded channels ($\tilde{4}$8-64) for this analysis, not just SMC channels. There are a total of 120 trials for each subject with 30 trials each corresponding to hand and tongue movements and interleaved with rest trials. We used TS learnt using finger flexion dataset to initialize the first layer of network for subjects in this new dataset to test inter dataset transfer learning (ITL-DNN).

3.2.12 Clustering brain areas

Functional mapping studies use behavioral labels and have some prior constraints on relevant cortical areas when looking for neural correlates. To evaluate if the representations learnt by AE can help identify functionally similar brain areas without any prior knowledge, we clustered electrode channels with the k-means algorithm based on their average activation across finger movement trials. We then validated the clustering results using two metrics.

- **Normalized Mutual Information (NMI)**: We compared our clusters to random clusters to validate that the clusters are informative. We have actual electrode locations (L_{act}) labeled by cortical regions based on anatomical boundaries. These regions do not form a ground truth for clustering, as electrodes

[1] Miller, Kai J., et al. "Power-law scaling in the brain surface electric potential." PLoS computational biology 5.12 (2009): e1000609.

from the same regions do not necessarily share the same physiological response patterns. However, physiological responses do tend to be more similar within regions than across regions and thus we expect a "good" clustering to share information with these regions. The random cluster labels (L_{rand}) were obtained by sampling from a uniform distribution of electrode location labels. NMI provides a metric to test if the cluster labels from AE (L_{ae}) are significantly different from L_{rand} and accounts for the varying number of clusters. NMI between two random variables X, Y is calculated using Eq 3.2, where $I(X,Y)$ denotes the mutual information between X, Y and $H(X)$ denotes the entropy of X. We calculated NMI between our cluster labels and L_{act} and compared this with the NMI between random cluster labels and L_{act}.

$$NMI(X,Y) = \frac{I(X;Y)}{[H(X)+H(Y)]/2} \tag{3.2}$$

- **Decoding accuracy**: We used groups of channels clustered together to decode finger flexions, using DNNs described earlier, to see if the task relevant electrodes tend to be clustered together.

3.3 Results

3.3.1 Decoding finger flexions

Classification performance comparing different feature extraction methods for the six subjects ('A'-'F') is shown in Figure 3.9. Results for excluded subjects are shown in Figure 3.14. Given the anatomical constraints of movements of ring and pinky fingers, their movements are highly correlated and hence; thus, we merged these two class labels into one class. Models using time-frequency features outperform models using SF, demonstrating the advantage of using features that leverage smaller frequency bins that span the entire broadband frequency range. Statistical significance was evaluated and details are provided in Tables 3.1 & 3.2. Subject F consistently had low decoding performance across all models; we suspect that this is due to insufficient ECoG grid coverage in SMC regions.

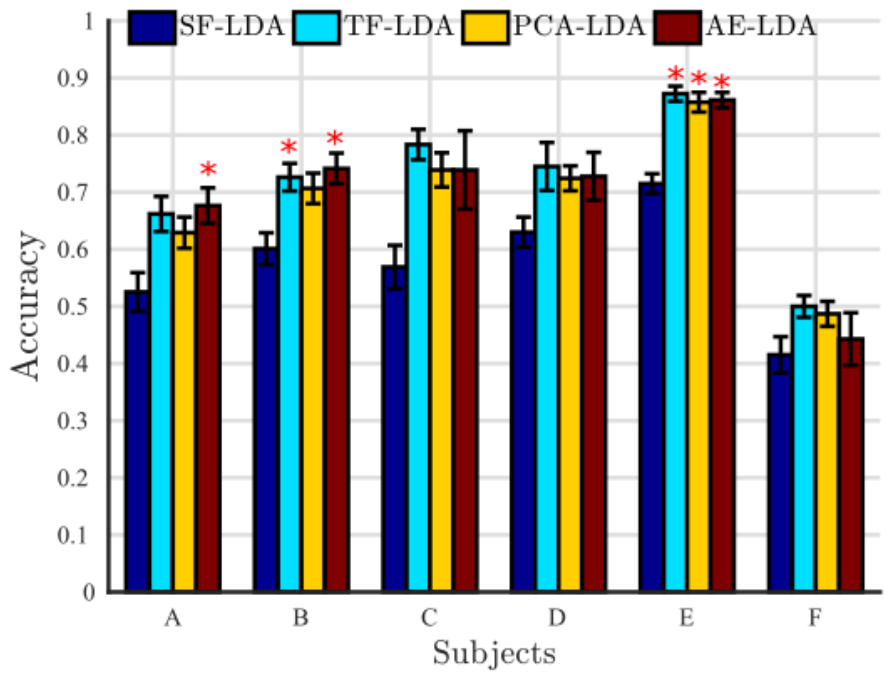

Figure 3.9: Comparing different feature extraction methods with LDA classifier. Mean accuracy with standard error bars over five cross validation folds for 4-class finger flexion classification is shown. Legend: SF-LDA: Spectral Features, TF-LDA: Features are estimated powers in Time-Frequency bins, PCA-LDA: Principal Component projections, AE-LDA: Autoencoder features. Chance level is 28% for random classifier. All the models performed above chance. Models with significantly higher performance compared to SF-LDA are indicated with red asterisks. ANOVA results comparing each condition to the other 3 are in 3.1.(P ¡ 0.05; multiple comparison with one-way ANOVA statistics.)

Although TF-LDA, AE-LDA and PCA-LDA have similar decoding performance against the original datasets collected in this study, autoencoders could provide more robust features under high noise conditions by incorporating a saturating non-linearity and learning sparse representations that are resilient to input dropout. Thus, in Figure 3.10 we measure the performance of these three decoding strategies with respect to increasing levels of additive spike noise. This analysis demonstrates that AE-LDA is more robust and can be advantages in noisy conditions which are common in EMUs.

LDA uses linear decision boundaries whereas DNN can model complex decision boundaries using non-linear activation functions. We compare how these different classifiers affect decoding performance in Figure 3.11. We show results from CK-DNN to show performance of spectral features with the advantage

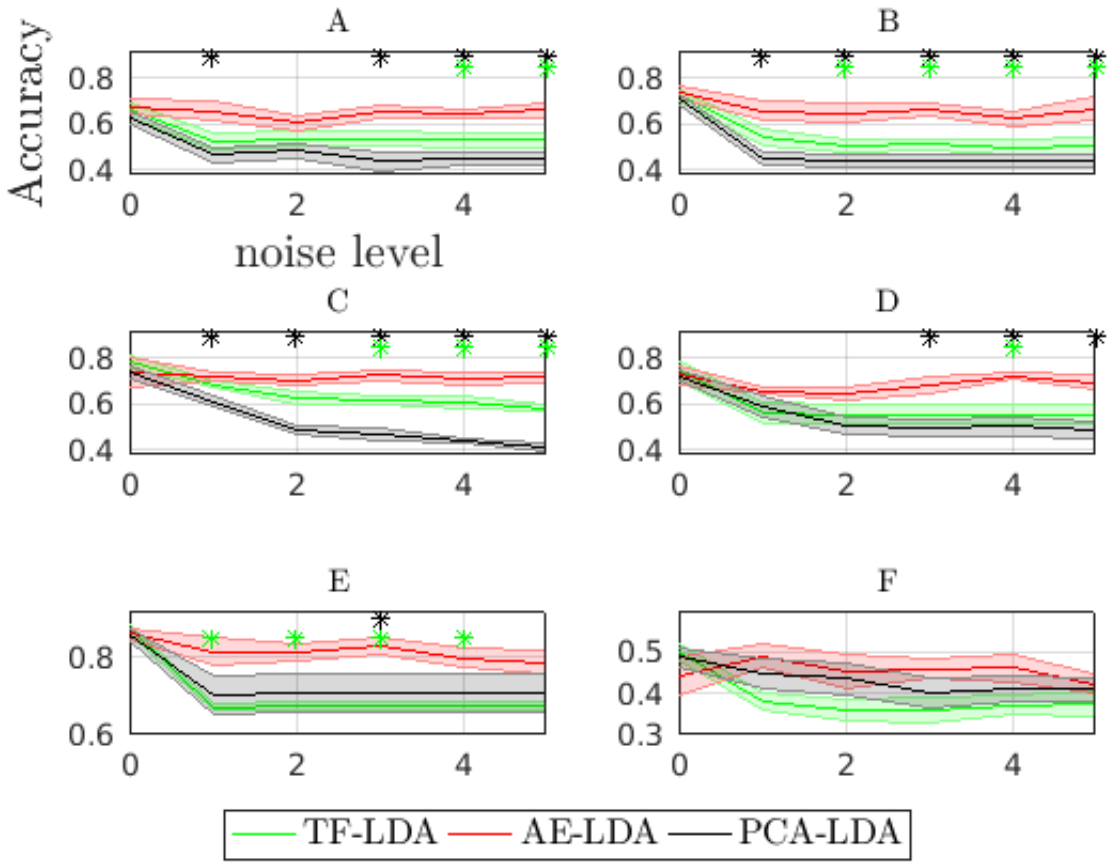

Figure 3.10: Noise levels which have a statistically significant difference between TF-LDA and AE-LDA are denoted with green asterisks and between the PCA-LDA and AE-LDA are denoted with black asterisks. ($p < 0.05$; multiple comparison with one-way ANOVA statistics.)

of having DNNs. While CK use powers in traditional frequency bands, the TS from autoencoders encode informative patterns learnt from provided data. By using TS from other subjects (TL-DNNs), performance is comparable to that achieved by using features obtained from same subject; this suggests that t-f patterns learnt across subjects hold generalizable information. Given the limited dataset sizes in ECoG studies, using features obtained via TL can be advantageous.

Although the performance of AE-DNN was observed to be on par with that of AE-LDA, DNNs could potentially benefit with increased trial count and have the potential to learn mappings for more complex tasks than LDA. In a more complex task, for example, behavioral labels could be arbitrary relative

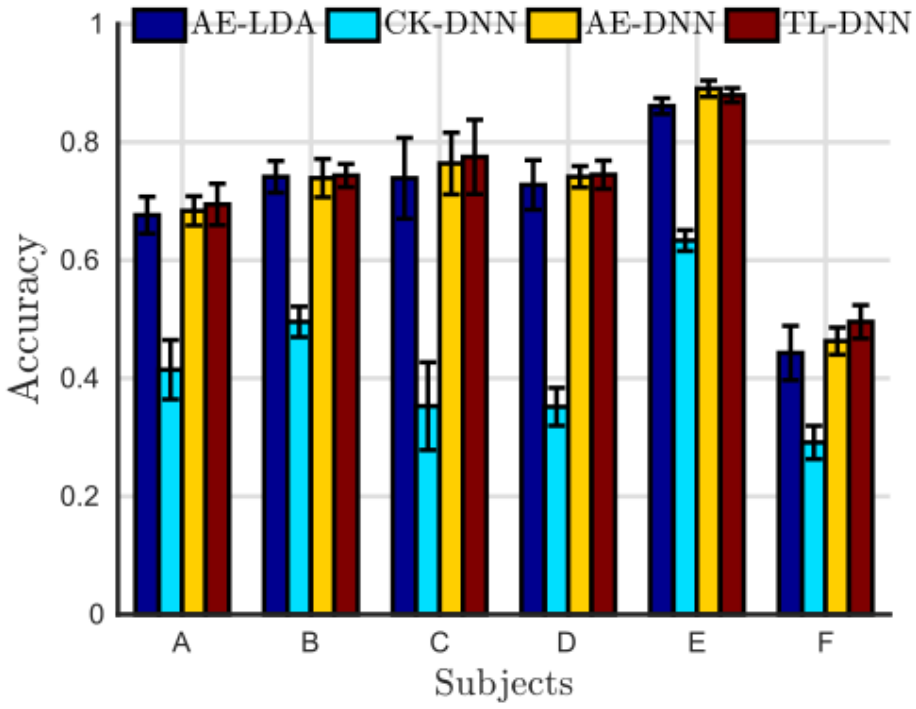

Figure 3.11: Comparing different classifiers. Mean accuracy with standard error bars over five cross validation folds for 4-class finger flexion classification is shown. Legend: AE-LDA: Autoencoder features with LDA classifier, CK-DNN: Custom Kernel features with Deep Neural Network classifier, AE-DNN: Subject specific autoencoder features with DNN, TL-DNN: Transfer Learning features from other subjects with DNN. AE-LDA, AE-DNN, TL-DNN show equivalent decoding performance. Chance level is 28% for random classifier. All the models, except for CK-DNN for subjects C, D and F, performed above chance.

to the true underlying behavioral intent and structure of neural activity. Conventional techniques like LDA typically assume that there is a prototypical pattern of neural activity for each behavioral label. The DNN, in contrast, has a structure that facilitates a more complex relationship between neural activity and behavioral label. To demonstrate this potential advantage, we consider a scenario in which two classes, thumb and index finger, which have distinct neural activity, are merged to create one label. Neural activity between thumb and index fingers is well distinguishable from one another compared to index and middle fingers. Therefore LDA would be able to find good decision boundaries for classification tasks when thumb, index, and middle fingers are labeled separately. However, when we merge thumb and index classes, effective classification might require a more complex mapping than possible via LDA. Figure 3.12 shows decoder performance for this modified classification task for AE-LDA and AE-DNN with 50 hidden nodes (arbitrary

parameter choice); Figure 3.15 shows the confusion matrices. DNNs fare well when decoding thumb/index vs middle finger demonstrating the comparative advantage DNNs have over LDA when labels are more abstract.

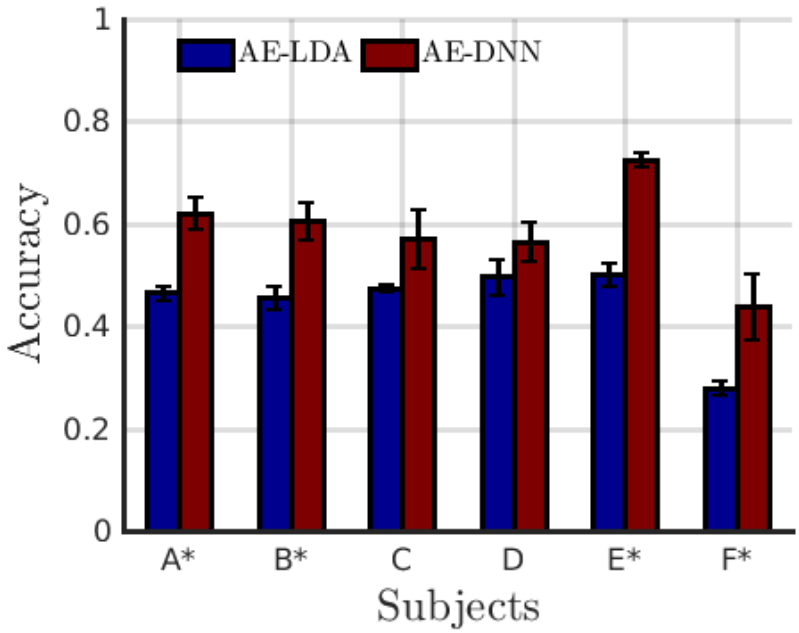

Figure 3.12: Comparing LDA and DNN when two classes (Thumb,Index fingers) are merged. Mean accuracies with standard error bars for 5 cross-validation folds are shown. Subjects indicated with asterisk have P¡0.05 with paired-sample t-Test.

The classification results for hand vs tongue dataset shown in Figure 3.13 suggest that the decoding method presented in this paper generalizes to other datasets. ITL-DNN performed on par with AE-DNN suggesting the plausibility of inter dataset transfer learning and AE based models outperformed SF-LDA despite the high dimensionality and small dataset size.

3.3.2 Significance tests

Significance tests were used to compare the different decoding methods for each subject. p-values from pairwise comparison of the subject's group means using Tukey's honestly significant difference criterion are shown in tables below. The statistics for multiple comparison tests are obtained from posthoc

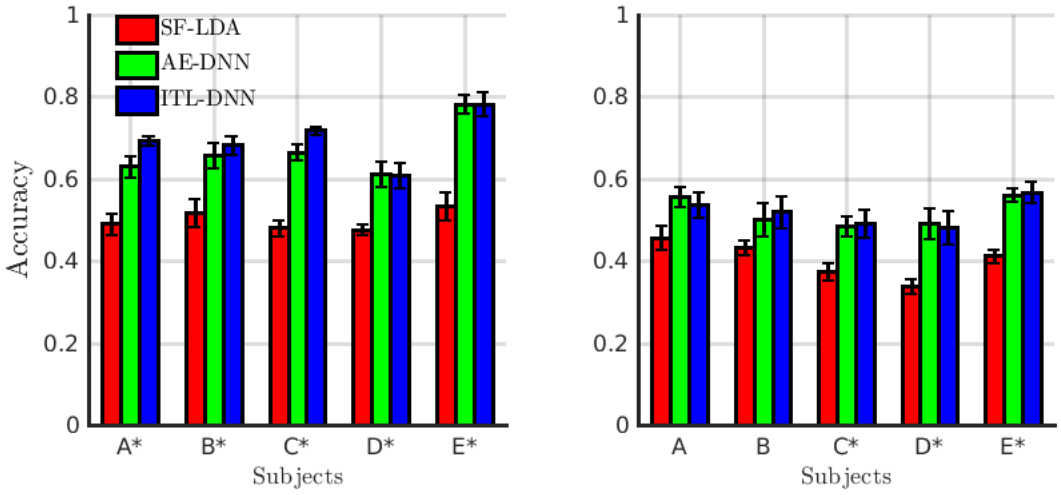

Figure 3.13: Subjects perform actual movement or kinesthetic imagery of hand and tongue. Mean accuracy with standard error bars over five cross validation folds for 3-class classification is shown. Chance level with random guessing is 37.5%. Actual movement decoding performance (left) is higher compared to kinesthetic imagery (right). For ITL-DNN, template spectrograms trained using finger flexion dataset were used in the first layer of DNN and for AE-DNN, template spectrograms were trained using respective subject's data. Subjects with significant difference ($p < 0.05$; multiple comparison with one-way ANOVA statistics.) between SF-LDA and AE-DNN are indicated with an asterisk.

testing after a one-way ANOVA. Tables 1 and 2 show p-values for 5 subjects A-F and G-I from multiple comparison tests to compare across models for each subject. p-values$<$ 0.05 are shown in bold face.

Table 3 shows the p-values for cluster decoding results. p-values of clusters with performance on par with SMC channels are shown in bold face.

3.3.3 Clustering results

Figure 3.16 demonstrate that NMI(L_{ae},L_{act}) and NMI(L_{rand},L_{act}) are significantly different (Results from t-test are in Table 3.3). Figure 3.18a is the brain maps when clustering electrode channels into 4 groups and Figure 3.18b quantifies classification performance using channels from each cluster. We observed that the cluster that significantly outperforms other clusters has performance comparable to that achieved when using SMC electrodes and tends to include electrodes from dorsal SMC, except for Subject F. Statistical comparisons of performance across clusters are detailed in Table 3.4 . This observation is consistent when

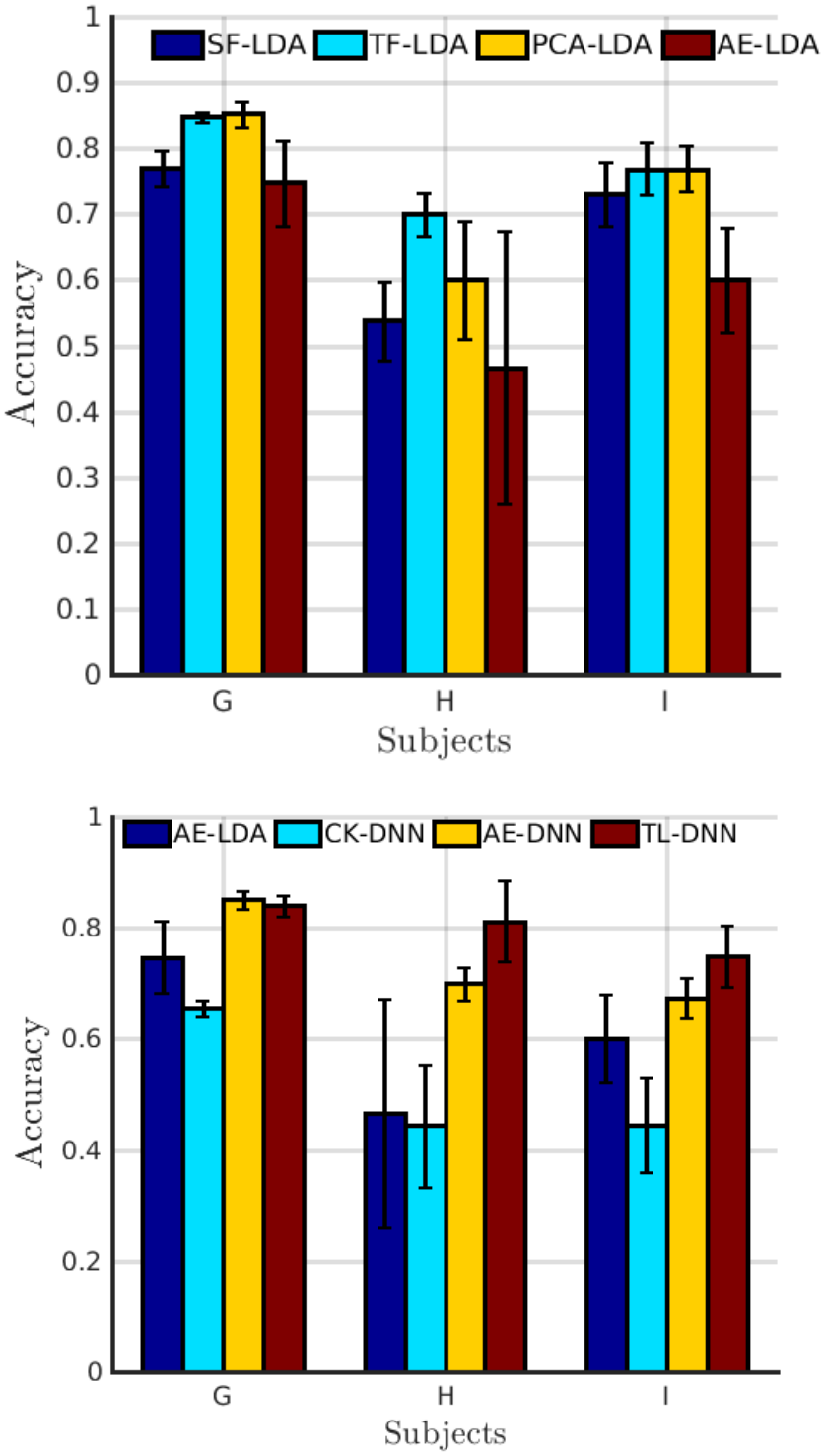

Figure 3.14: Mean accuracy with standard error bars over five cross validation folds for 4-class finger flexion classification comparing different feature extraction methods (top) and different classifiers (bottom). Chance level is 28% for random classifier.

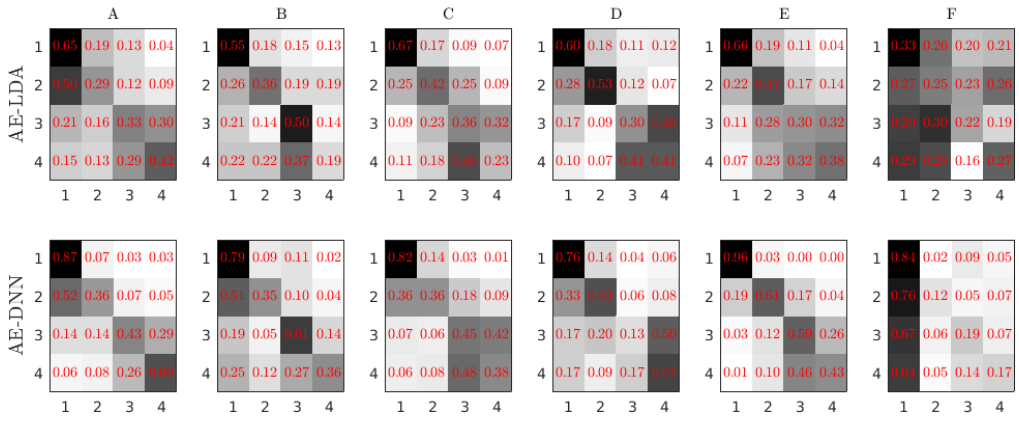

Figure 3.15: Sum of confusion matrices from five cross-validation folds of subjects A-E when Thumb and Index fingers are merged to create one class. The four labels are 1:'Thumb/Index', 2:'Middle',3:'Ring' and 4:'Pinky' fingers. Confusion matrices show true labels on vertical axis and predicted labels on horizontal axis.

varying the number of clusters. Further, we performed the same clustering experiment with only SMC electrodes and observed that electrodes in dorsal SMC and ventral SMC are almost always clustered into two groups (see Figure 3.17) and the electrode cluster with higher classification performance is in dorsal SMC. This is in agreement with related studies [15, 43] that document distinct dorsal SMC activity during finger flexions. This suggests that the proposed clustering based dimensionality reduction method can be used to identify task relevant channels in an unsupervised fashion.

3.4 Discussion

During motor activity, significant increases in ECoG signal power occurs mostly in high frequency bands. However, it is difficult to estimate the exact frequency ranges and the timing of increased activity for different electrode channels. We exploit the temporal and frequency information in the data by restructuring signals into a t-f format. If we consider the result from Figure 4, we see that TF-LDA can outperform SF-LDA, suggesting that a finer frequency resolution is advantageous in this task. However, the TF-LDA approach may lead to poor generalization as the number of features grows to the product of number of

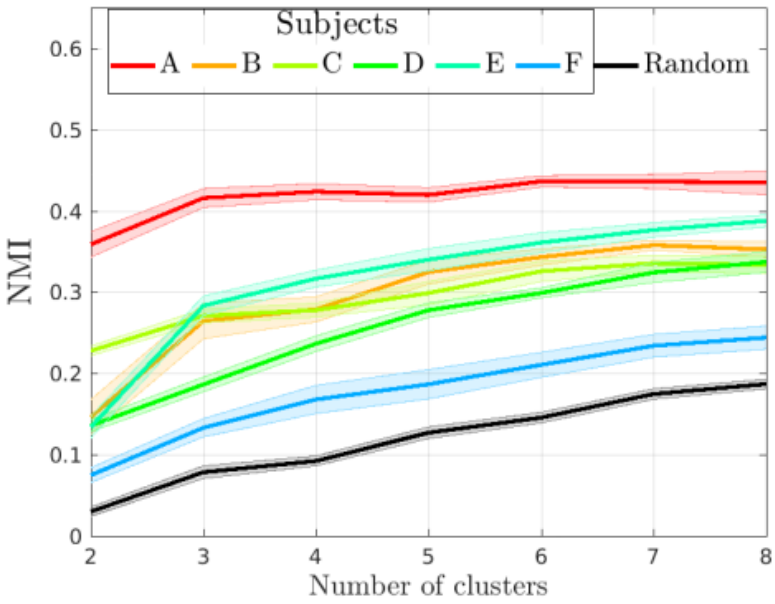

Figure 3.16: NMI is calculated for 10 random initializations and the mean values with standard error bars are shown.

channels, frequency bins, and time bins. In the current study, the ratio of TF-LDA features to trials was approximately 15 to 1 for most subjects. We noticed that regularization of TF-LDA tended to reduce the effective parameterization of TF-LDA by weighting off-diagonal terms of the covariance matrix in the LDA model close to zero. The resulting model is effectively a Naive Bayes model which assumes all features (i.e. all t-f entries) are independent given the finger being flexed. Thus, we compare these results to PCA-LDA and AE-LDA; AE and PCA (which is the special linear case of AE) can efficiently aggregate data from all frequency bands and time bins based upon underlying regularities (i.e. correlations between t-f entries) in the dataset. AE and PCA allow the models to move away from using fixed frequency band spectral features selected based upon previous literature while controlling the number of features used by the decoding model. The comparable performance for TF-LDA, PCA-LDA, and AE-LDA in Figure 4 suggests that this additional structure can be imposed upon t-f based features without a loss in performance.

Figure 3.19 shows sample t-f responses of an SMC electrode. Since for the majority of electrodes' t-f responses vary in power for specific high frequency bands in narrow time bins, the autoencoder preserves

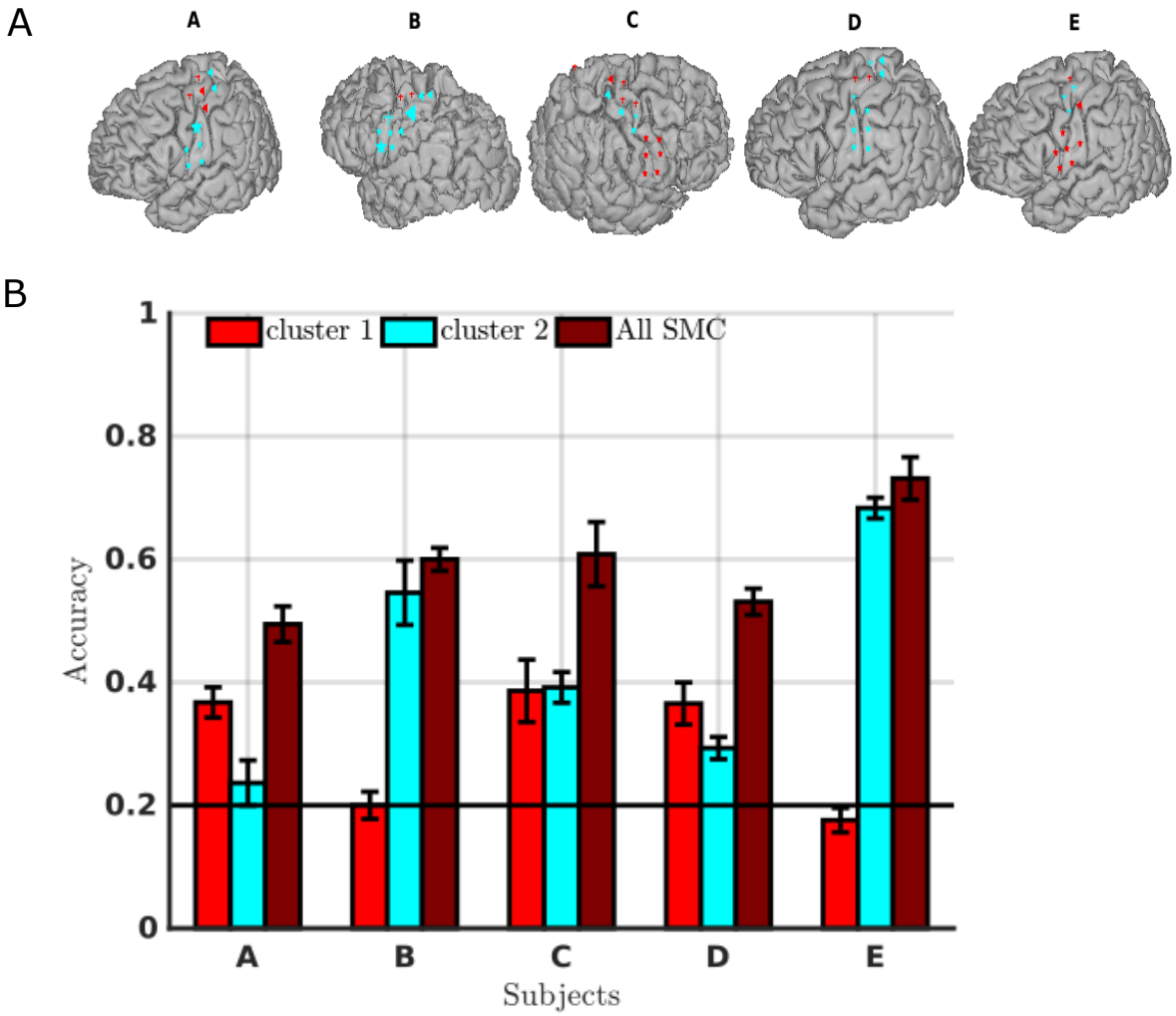

Figure 3.17: A shows the brain maps with electrode clusters. B shows 5 class classification performance with AE-DNN with $h_1 = 60$ and $h_2 = 10$.

this information by learning kernels that depict bandpass filters at different time bins as shown in Figure 3a. These template spectrograms generated using AE when used in a convolutional neural network framework serve as filters learnt in a data-driven approach. The AE used in this work uses a non-linear transform, the satlin function. While the addition of this non-linearity does not convey a measurable benefit relative to PCA for the dataset, we demonstrate in Figure 5 how this approach can increase model robustness in higher

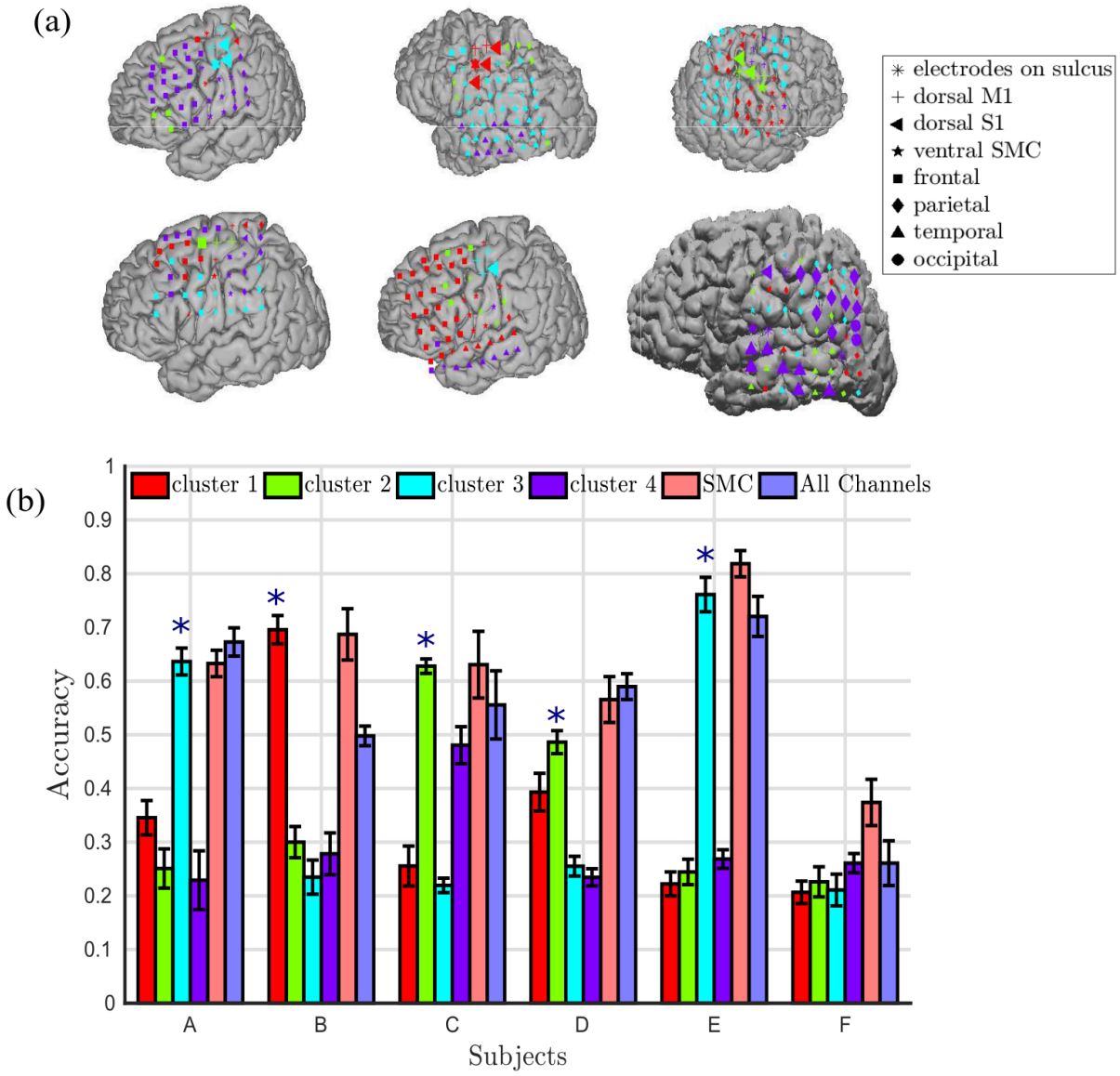

Figure 3.18: (a): Electrode channels projected onto brain images for five subjects. Colors correspond to different clusters from k-means algorithm. Shapes correspond to original channel locations as obtained from brain images. Note that for subjects A, D and E, electrodes are projected on a standard template and for subjects B, C and F, subjects' original brain maps are used. (b): Decoding performance results using electrode channels from each cluster, SMC channels, and all channels. Colors of cluster bars match with the cluster colors in Figure 3.18a. The figure shows 4 class classification performance with AE-DNN with fixed (non-optimized) hyperparamter values, $h_1 = 60$ and $h_2 = 10$. Blue asterisks above the bar indicates cluster decoding performance on par with SMC channels decoding performance.

noise conditions.

DL methods are successfully being employed across various domains such as computer vision,

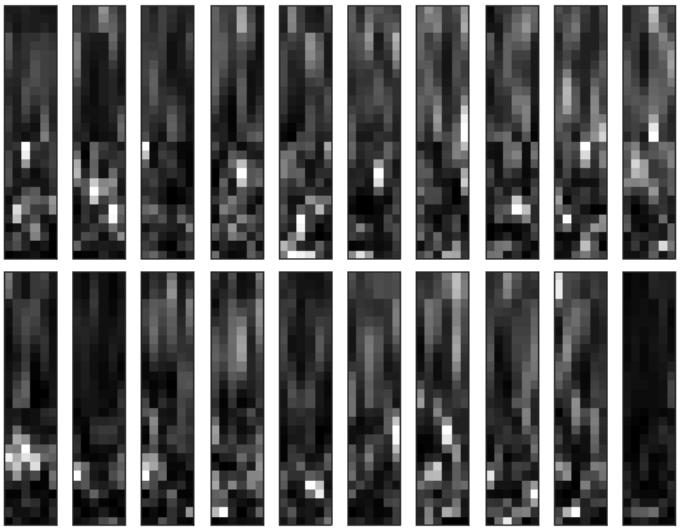

Figure 3.19: We can notice increase in power in specific frequency ranges at varying time bins. Autoencoder kernels shown in Figure 3 in main manuscript are learnt and appear to summarize these t-f responses of this nature.

speech recognition, and natural language processing [44–46]. However, the major impediment for using DL in neurophysiology studies is the lack of sufficient data to train large scale networks with millions of parameters. The proposed DNN architecture has a modest parameter count (of the order of tens of thousands) which can be advantageous for neural decoding when data size is limited. One downside of using DNN methods is the additional effort required for hyperparameter tuning. In our analyses, we used a small grid search on the number of hidden nodes and fixed the learning rate constant. However, one could complete a finer grid search and additional hyperparamater tuning to potentially improve decoder performance at additional computational expense.

Figure 3.12 demonstrates that DNNs can be advantageous in learning complex mappings under ambiguous or abstract behavioral conditions. This proof-of-concept analysis demonstrates the potential advantage of employing DNN architectures in situations in which behaviors are ambiguous or not well defined. Another advantage of using DNN architectures is that they provide the opportunity to integrate feature extraction methods like AE. By adding the trained AE encoder as an input layer of a DNN classifier, we also

have the opportunity to optimize AE weights to better suit the classification problem by backpropagating errors through the single aggregate DNN during training. This could be especially useful in a transfer learning setting, where AE weights are first learnt in an unsupervised fashion and then fine-tuned in combination with DNNs for the classification task at hand.

With advancements in high density ECoG recording technology [47–50], neurophysiology datasets are likely to see an orders of magnitude increase in data dimensionality. To begin to understand these complex datasets, we must leverage the availability of long duration EMU recordings that can span a few days. However, excluding the recordings collected during structured behavioral tasks, which is of the order of few minutes, a vast portion of the ECoG data is unlabeled or coarsely labeled. The TL-DNN results suggest that a library of t-f patterns learnt across such unlabeled datasets can be used to initiate deep convolutional networks in data limited studies allowing for knowledge transfer.

Surgical procedure for treating epilepsy involves electrically stimulating cortical surfaces through electrodes on grids. With emerging research in high density ECoG, electrode count on grids has increased dramatically. The unsupervised clustering based approach can potentially help clinical functional mapping by reducing the cortical areas of interest to a few localized clusters. While the clustering analysis presented in Figures 9,10 can be done with SF or PCA features, we used the t-f representations from autoencoder, since Figures 4 and 5 suggest that the autoencoder based features are more robust than PCA based features and are more informative than spectral features.

This chapter is a reprint of the material as it appears in Journal of Neural Engineering (2018), "Autoencoders for learning template spectrograms in electrocorticographic signals." Pailla, Tejaswy, Kai Joshua Miller, and Vikash Gilja. The dissertation author is the primary investigator and author of this paper.

Table 3.1: p-values from multiple comparison tests comparing different feature extraction methods and decoding methods for subjects A-F

Models		Subjects					
		A	B	C	D	E	F
SF-LDA	TF-LDA	0.106	**0.0338**	0.103	0.153	**1.08e-06**	0.418
	PCA-LDA	0.354	0.106	0.303	0.346	**6.18e-06**	0.609
	AE-LDA	0.0555	**0.0127**	0.303	0.306	**3.96e-06**	0.994
	CK-DNN	0.285	0.107	0.0978	**1.26e-05**	**0.0112**	0.0778
	AE-DNN	**0.0396**	**0.0147**	0.17	0.177	**1.58e-07**	0.908
	TL-DNN	**0.0233**	**0.011**	0.128	0.153	**4.72e-07**	0.48
TF-LDA	PCA-LDA	0.993	0.998	0.997	0.999	0.992	1
	AE-LDA	1	1	0.997	1	0.998	0.813
	CK-DNN	**0.000369**	**2.26e-05**	**7.69e-05**	**5.47e-08**	**3.72e-08**	**0.000495**
	AE-DNN	0.999	1	1	1	0.975	0.972
	TL-DNN	0.993	0.999	1	1	1	1
PCA-LDA	AE-LDA	0.955	0.964	1	1	1	0.935
	CK-DNN	**0.00221**	**9.06e-05**	**0.00037**	**9.12e-08**	**3.78e-08**	**0.00113**
	AE-DNN	0.914	0.974	1	1	0.71	0.997
	TL-DNN	0.821	0.952	0.999	0.999	0.941	1
AE-LDA	CK-DNN	**0.000165**	**7.78e-06**	**0.00037**	**8.19e-08**	**3.76e-08**	**0.002**
	AE-DNN	1	1	1	1	0.806	0.999
	TL-DNN	1	1	0.999	1	0.975	0.862
CK-DNN	AE-DNN	**0.000111**	**9.05e-06**	**0.000153**	**5.83e-08**	**3.71e-08**	**0.00506**
	TL-DNN	**6.06e-05**	**6.69e-06**	**0.000103**	**5.47e-08**	**3.72e-08**	**0.0123**
AE-DNN	TL-DNN	1	1	1	1	0.998	0.985

Table 3.2: p-values for excluded subjects G-I.

Models		Subjects		
		G	H	I
SF-LDA	TF-LDA	0.549	0.917	0.999
	PCA-LDA	0.471	0.999	0.999
	AE-LDA	0.998	0.999	0.675
	CK-DNN	0.146	0.994	**0.0215**
	AE-DNN	0.497	0.917	0.992
	TL-DNN	0.655	0.5	1
TF-LDA	PCA-LDA	1	0.992	1
	AE-LDA	0.265	0.678	0.394
	CK-DNN	**0.00199**	0.58	**0.00688**
	AE-DNN	1	1	0.906
	TL-DNN	1	0.986	1
PCA-LDA	0.935	0.213	0.966	0.394
	CK-DNN	**0.00143**	0.93	**0.00688**
	AE-DNN	1	0.992	0.906
	TL-DNN	1	0.762	1
AE-LDA	CK-DNN	0.356	1	0.484
	AE-DNN	0.229	0.678	0.966
	TL-DNN	0.346	0.243	0.531
CK-DNN	AE-DNN	**0.0016**	0.58	0.102
	TL-DNN	**0.00309**	0.183	**0.0123**
AE-DNN	TL-DNN	1	0.986	0.966

Table 3.3: p-values for NMI comparisons shown in Fig 6

Subjects	Number of Clusters						
	2	3	4	5	6	7	8
A	3.7336e-08	2.0011e-11	2.0101e-09	8.0431e-08	1.1144e-06	5.262e-07	3.3203e-07
B	0.00049218	1.6744e-06	8.4533e-06	3.2309e-05	5.4786e-05	0.00025408	0.00075218
C	1.1852e-06	1.6533e-05	2.3306e-06	9.4869e-05	6.7712e-06	1.8027e-05	5.1248e-05
D	0.0006323	0.038785	0.0042044	0.026984	0.035542	0.010361	0.054018
E	0.00057913	1.1409e-06	1.0596e-06	4.5653e-07	5.3455e-08	4.3602e-10	3.3207e-09
F	0.0006	0.0011e-03	0.0011e-03	0.005e-03	1.3e0-9	1.2e-09	3.32e-09

Table 3.4: p-values for clustering results. (Ci denotes Cluster i)

Models		Subjects					
		A	B	C	D	E	F
C1	2	0.41571	2.1512e-07	2.8206e-05	0.21147	0.99135	0.9977
	3	5.7835e-05	3.1934e-08	0.98992	0.020581	2.0693e-08	1
	4	0.20957	9.3885e-08	0.01131	0.0059917	0.82704	0.8
	SMC	6.9223e-05	**0.99997**	2.5236e-05	0.0025648	2.0688e-08	0.01
	All Chan	9.8504e-06	0.0039328	0.00053721	0.00056861	2.0719e-08	0.82
C2	C3	6.7442e-07	0.73773	6.7598e-06	6.6389e-05	2.0702e-08	1
	C4	0.99757	0.99709	0.17854	1.8748e-05	0.98757	0.96
	SMC	7.9123e-07	3.1028e-07	**1**	**0.36759**	2.0689e-08	0.03
	All Chan	1.5192e-07	0.0039328	0.83113	0.13137	2.077e-08	0.96
C3	C4	2.6806e-07	0.93735	0.002654	0.99467	2.0727e-08	0.86
	SMC	**1**	3.6906e-08	6.0671e-06	6.2978e-07	**0.66615**	0.01
	All Chan	0.97503	0.00013105	0.00012183	1.8151e-07	0.88987	0.86
C4	SMC	3.1104e-07	1.2853e-07	**0.16418**	2.1463e-07	2.0691e-08	0.15
	All Chan	7.2473e-08	0.001272	0.80847	7.4041e-08	2.0929e-08	1.00
SMC	All Chan	0.96254	0.0061398	0.80847	0.98918	0.14423	0.15

References

[1] Nathan E Crone, Alon Sinai, and Anna Korzeniewska. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Progress in brain research*, 159:275–295, 2006.

[2] Kai J Miller, Pradeep Shenoy, John W Miller, Rajesh PN Rao, Jeffrey G Ojemann, et al. Real-time functional brain mapping using electrocorticography. *Neuroimage*, 37(2):504–507, 2007.

[3] Eric C Leuthardt, Gerwin Schalk, Jonathan R Wolpaw, Jeffrey G Ojemann, and Daniel W Moran. A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, 1(2):63, 2004.

[4] Nathan E Crone, Diana L Miglioretti, Barry Gordon, Jeffrey M Sieracki, Michael T Wilson, Sumio Uematsu, and Ronald P Lesser. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. i. alpha and beta event-related desynchronization. *Brain*, 121(12):2271–2299, 1998.

[5] Nathan E Crone, Diana L Miglioretti, Barry Gordon, and Ronald P Lesser. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. ii. event-related synchronization in the gamma band. *Brain*, 121(12):2301–2315, 1998.

[6] Kristofer E Bouchard, Nima Mesgarani, Keith Johnson, and Edward F Chang. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441):327–332, 2013.

[7] Xiaomei Pei, Eric C Leuthardt, Charles M Gaona, Peter Brunner, Jonathan R Wolpaw, and Gerwin Schalk. Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *Neuroimage*, 54(4):2960–2972, 2011.

[8] Camilo Toro, Günther Deuschl, Robert Thatcher, Susumu Sato, Conrad Kufta, and Mark Hallett. Event-related desynchronization and movement-related cortical potentials on the ecog and eeg. *Electroencephalography and clinical neurophysiology/evoked potentials section*, 93(5):380–389, 1994.

[9] Philip R Kennedy, M Todd Kirby, Melody M Moore, Brandon King, and Adon Mallory. Computer control using human intracortical local field potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(3):339–344, 2004.

[10] Takufumi Yanagisawa, Masayuki Hirata, Youichi Saitoh, Haruhiko Kishima, Kojiro Matsushita, Tetsu Goto, Ryohei Fukuma, Hiroshi Yokoi, Yukiyasu Kamitani, and Toshiki Yoshimine. Electrocorticographic control of a prosthetic arm in paralyzed patients. *Annals of neurology*, 71(3):353–361, 2012.

[11] Guy Hotson, David P McMullen, Matthew S Fifer, Matthew S Johannes, Kapil D Katyal, Matthew P Para, Robert Armiger, William S Anderson, Nitish V Thakor, Brock A Wester, et al. Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject. *Journal of neural engineering*, 13(2):026017, 2016.

[12] Alan D Degenhart, Shivayogi V Hiremath, Ying Yang, Stephen Foldes, Jennifer L Collinger, Michael Boninger, Elizabeth C Tyler-Kabara, and Wei Wang. Remapping cortical modulation for electrocorticographic brain–computer interfaces: a somatotopy-based approach in individuals with upper-limb paralysis. *Journal of neural engineering*, 15(2):026021, 2018.

[13] Mariska J Vansteensel, Elmar GM Pels, Martin G Bleichner, Mariana P Branco, Timothy Denison, Zachary V Freudenburg, Peter Gosselaar, Sacha Leinders, Thomas H Ottens, Max A Van Den Boom,

et al. Fully implanted brain–computer interface in a locked-in patient with als. *New England Journal of Medicine*, 375(21):2060–2066, 2016.

[14] Kai J Miller, Eric C Leuthardt, Gerwin Schalk, Rajesh PN Rao, Nicholas R Anderson, Daniel W Moran, John W Miller, and Jeffrey G Ojemann. Spectral changes in cortical surface potentials during motor movement. *The Journal of neuroscience*, 27(9):2424–2432, 2007.

[15] KJ Miller, S Zanos, EE Fetz, M Den Nijs, and JG Ojemann. Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans. *Journal of Neuroscience*, 29(10):3132–3137, 2009.

[16] Eric C Leuthardt, Kai J Miller, Gerwin Schalk, Rajesh PN Rao, and Jeffrey G Ojemann. Electrocorticography-based brain computer interface-the seattle experience. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):194–198, 2006.

[17] Jeremiah D Wander, Timothy Blakely, Kai J Miller, Kurt E Weaver, Lise A Johnson, Jared D Olson, Eberhard E Fetz, Rajesh PN Rao, and Jeffrey G Ojemann. Distributed cortical adaptation during learning of a brain–computer interface task. *Proceedings of the National Academy of Sciences*, 110(26):10818–10823, 2013.

[18] Sergey M Plis, Devon R Hjelm, Ruslan Salakhutdinov, Elena A Allen, Henry J Bockholt, Jeffrey D Long, Hans J Johnson, Jane S Paulsen, Jessica A Turner, and Vince D Calhoun. Deep learning for neuroimaging: a validation study. *Frontiers in neuroscience*, 8:229, 2014.

[19] Hubert Cecotti and Axel Graser. Convolutional neural networks for p300 detection with application to brain-computer interfaces. *IEEE transactions on pattern analysis and machine intelligence*, 33(3):433–445, 2011.

[20] Piotr Mirowski, Deepak Madhavan, Yann LeCun, and Ruben Kuzniecky. Classification of patterns of eeg synchronization for seizure prediction. *Clinical neurophysiology*, 120(11):1927–1940, 2009.

[21] DF Wulsin, JR Gupta, R Mani, JA Blanco, and B Litt. Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement. *Journal of neural engineering*, 8(3):036015, 2011.

[22] Martin Längkvist, Lars Karlsson, and Amy Loutfi. Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems*, 2012:5, 2012.

[23] JT Turner, Adam Page, Tinoosh Mohsenin, and Tim Oates. Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection. In *2014 AAAI Spring Symposium Series*, 2014.

[24] Zachary V Freudenburg, Nicolas F Ramsey, Mark Wronkiewicz, William D Smart, Robert Pless, and Eric C Leuthardt. Real-time naive learning of neural correlates in ecog electrophysiology. *International Journal of Machine Learning and Computing*, 1(3):269, 2011.

[25] Zuoguan Wang, Siwei Lyu, Gerwin Schalk, and Qiang Ji. Deep feature learning using target priors with applications in ecog signal decoding for bci. In *IJCAI*, 2013.

[26] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *International Conference on Artificial Neural Networks*, pages 52–59. Springer, 2011.

[27] JC Sanchez, JC Principe, JM Carmena, Mikhail A Lebedev, and MAL Nicolelis. Simultaneus prediction of four kinematic variables for a brain-machine interface using a single recurrent neural network. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 5321–5324. IEEE, 2004.

[28] Yadunandana N Rao, S-P Kim, Justin C Sanchez, Deniz Erdogmus, Jose C Principe, Jose M Carmena, Mikhail A Lebedev, and Miguel A Nicolelis. Learning mappings in brain machine interfaces with echo state networks. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 5, pages v–233. IEEE, 2005.

[29] Aysegul Gunduz, Mustafa C Ozturk, Justin C Sanchez, and Jose C Principe. Echo state networks for motor control of human ecog neuroprosthetics. In *2007 3rd International IEEE/EMBS Conference on Neural Engineering*, pages 514–517. IEEE, 2007.

[30] David Sussillo, Paul Nuyujukian, Joline M Fan, Jonathan C Kao, Sergey D Stavisky, Stephen Ryu, and Krishna Shenoy. A recurrent neural network for closed-loop intracortical brain–machine interface decoders. *Journal of neural engineering*, 9(2):026027, 2012.

[31] Pouya Bashivan, Irina Rish, Mohammed Yeasin, and Noel Codella. Learning representations from eeg with deep recurrent-convolutional neural networks. *ICLR 2016,arXiv preprint arXiv:1511.06448*, 2016.

[32] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1, 1988.

[33] Kai J Miller, Dora Hermes, Christopher J Honey, Adam O Hebb, Nick F Ramsey, Robert T Knight, Jeffrey G Ojemann, and Eberhard E Fetz. Human motor cortical activity is selectively phase-entrained on underlying rhythms. *PLoS Comput Biol*, 8(9):e1002655, 2012.

[34] KJ Miller and JG Ojemann. A library of human electrocorticographic data and analyses. *Program No. 469.08. 2016 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2016. Online*.

[35] Dora Hermes, Kai J Miller, Herke Jan Noordmans, Mariska J Vansteensel, and Nick F Ramsey. Automated electrocorticographic electrode localization on individually rendered brain surfaces. *Journal of neuroscience methods*, 185(2):293–298, 2010.

[36] David DeMers and Garrison W Cottrell. Non-linear dimensionality reduction. In *Advances in neural information processing systems*, pages 580–587, 1993.

[37] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[38] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103. ACM, 2008.

[39] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

[40] A. Vedaldi and K. Lenc. Matconvnet – convolutional neural networks for matlab. In *Proceeding of the ACM Int. Conf. on Multimedia*, 2015.

[41] Pradeep Shenoy, Kai J Miller, Jeffrey G Ojemann, and Rajesh PN Rao. Generalized features for electrocorticographic bcis. *Biomedical Engineering, IEEE Transactions on*, 55(1):273–280, 2008.

[42] Kai J Miller, Gerwin Schalk, Eberhard E Fetz, Marcel Den Nijs, Jeffrey G Ojemann, and Rajesh PN Rao. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, 107(9):4430–4435, 2010.

[43] KJ Kubanek, Miller, JG Ojemann, JR Wolpaw, and G Schalk. Decoding flexion of individual fingers using electrocorticographic signals in humans. *Journal of neural engineering*, 6(6):066001, 2009.

[44] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[45] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[46] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167. ACM, 2008.

[47] Jonathan Viventi, Dae-Hyeong Kim, Leif Vigeland, Eric S Frechette, Justin A Blanco, Yun-Soung Kim, Andrew E Avrin, Vineet R Tiruvadi, Suk-Won Hwang, Ann C Vanleer, et al. Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity in vivo. *Nature neuroscience*, 14(12):1599–1605, 2011.

[48] Edward F Chang. Towards large-scale, human-based, mesoscopic neurotechnologies. *Neuron*, 86(1):68–78, 2015.

[49] John Hermiz, Nicholas Rogers, Erik Kaestner, Mehran Ganji, Daniel R Cleary, Bob S Carter, David Barba, Shadi A Dayeh, Eric Halgren, and Vikash Gilja. Sub-millimeter ecog pitch in human enables higher fidelity cognitive neural state estimation. *NeuroImage*, 176:454–464, 2018.

[50] Mehran Ganji, Erik Kaestner, John Hermiz, Nick Rogers, Atsunori Tanaka, Daniel Cleary, Sang Heon Lee, Jospeh Snider, Milan Halgren, Garth Rees Cosgrove, et al. Development and translation of pedot: Pss microelectrodes for intraoperative monitoring. *Advanced Functional Materials*, 28(12):1700232, 2018.

Chapter 4

Cross-subject learning for decoding ECoG

4.1 Introduction

Prosthetic devices or brain-machine interfaces (BMIs) aim to restore the lost mobility in individuals paralyzed due to stroke, spinal cord injury or neurodegenerative disorders like amyotropic lateral sclerosis. BMIs decode intent from neural signals by mapping activity of neural ensembles to behaviors. Invasive BMIs with intracortical signals, recorded using Utah arrays [1], as control signals have been shown to successfully decode volitional movements and enabled users, paralyzed from neck down, to perform tasks like cursor control, reaching and grasping [2–4]. Electrocorticography (ECoG) is another alternative recording modality wherein neural activity is recorded from the cortical surface. Although the current clinical ECoG electrodes record signals from a bigger ensemble of neurons compared to intracortical electrodes, successful decoding of volitional motor activity from ECoG has been shown to be possible [5, 6]. The viability of ECoG as a neural acquisition modality for prosthetic control has been an area of active research for the past two decades [7–9]. The feasibility of ECoG based BMIs for paralyzed individuals has been demonstrated in [10–12]. ECOG based BMIs have the potential of long term stability in performance [13]

Training a decoder for BMI involves mapping spatio-temporal patterns in neural data to behaviors. Transfer of BMI decoders across multiple users or across session for a given user is a challenging problem.

Performance of decoder might vary across sessions for a given user. This could be due to change in cortical activity as user's familiarity with task changes [14], statistical variability in neural recordings due to changes in impedance or movements of electrodes, and hardware failure leading to loss of recordings from electrodes [15]. Decoders are typically recalibrated prior to session to ensure the neural to behavioral mapping is finetuned to the BMI user's neural activity for that particular session [16]. An alternative approach to improving BMI decoder's across-session stability is to use domain adaptation techniques to align the latent spaces of features across sessions to that of the first session on which the BMI decoder is trained [17]. In [18], the authors describe the idea of aligning distributions of low-dimensional projections of cortical activity with distribution of decoder outputs, thus relying on statistics of movements and neural activity to build decoders.

In this paper, we explore the idea of cross-subject learning (CL) of ECoG based BMI decoders, i.e. training a decoder using data from a set of users and deploying it for a new user. CL is a well studied problem in electroencephalography (EEG) based BMIs [19–23]. EEG caps have standard system for the location of scalp electrodes, making it easier to standardize datasets across users. However, CL is not straightforward for ECoG based BMIs. Majority of ECoG based studies rely on datasets collected from epilepsy monitoring units (EMUs). Individuals with chronic epilepsy undergo surgery for ECoG implantation to identify seizure foci. These individuals volunteer to participate in experiments while in clinical care post electrode implantation. The electrode grid placement is guided by individual subject's clinical requirements. This leads to variability in recorded signals across subjects in addition to to anatomical differences of cortex across users.

We address the challenge of cross-subject learning for ECoG based BMIs and demonstrate our results using a finger flexion dataset for both discrete and continuous decoding. We hypothesize that, given a set of users performing a behaviorally similar task (e.g. flexing specific finger), we can learn a common latent space from the neural activity corresponding to the common behavioral activity. We can then learn a common decoder across subjects that maps the dynamics from this latent space to the behavioral output.

We will discuss the effect of leveraging information across multiple subjects' data doing similar behavioral tasks on BMI decoder performance.

4.2 Methods

4.2.1 Dataset

We used the finger flexion dataset that originally appeared in [24] and was publicly released in [25]. Subjects with ECoG grids implanted for epilepsy treatment volunteered for the experiment. Subjects performed a finger flexion task wherein they were instructed to move specific fingers of the hand contralateral to the implant in response to visual cues. We use data from 5 subjects ('A'-'E') and a comprehensive description of dataset and subject exclusion criteria is described in [26]. ECoG signals were amplified, bandpass filtered between 0.15 and 200 Hz, digitized at 1000 Hz and stored along with the time synchronized digitized flexion samples for all fingers. All the subjects participated in a purely voluntary manner, after providing informed written consent, under experimental protocols approved by the Institutional Review Board of the University of Washington (#12193). All patient data were anonymized according to IRB protocol, in accordance with HIPAA mandate.

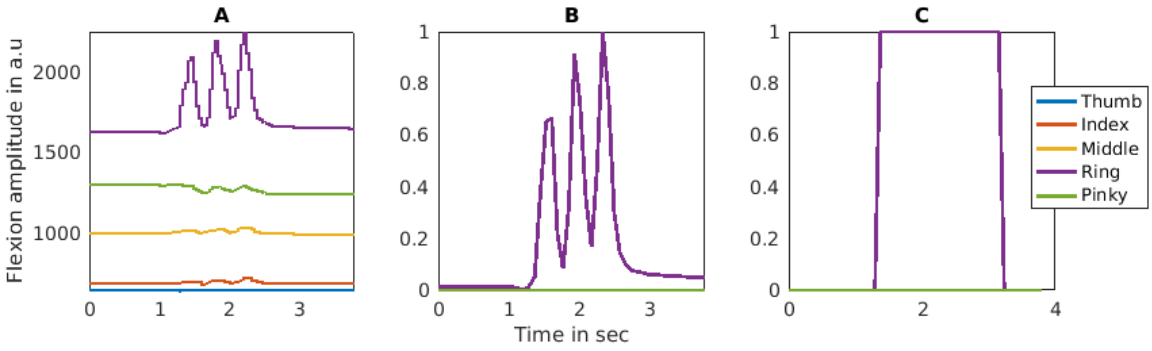

Figure 4.1: Preprocessing flexion trajectories from data glove. **A**: Raw flexion trajectory from data glove. **B** Cleaned trajectory for continuous decoding. **C**: Thresholded amplitude for discrete classification

4.2.2 Preprocessing data

Finger flexion timings for all subjects are manually annotated. Trial length for both continuous and discrete decoding is 4 sec. ECoG data and flexion trajectories from glove are preprocessed and downsampled to 12.5 Hz, thus resulting in 80 ms time step prediction, as described below.

Glove data: Flexion samples from data glove are smoothened with a 80 ms moving average window and detrended. They are then downsampled to 12.5 Hz and zscored. For continuous regression, flexion trajectories of the moving finger are normalized to the range [0, 1]. For discrete decoding, binary labels are assigned with unit amplitude during the movement period for the flexed finger (See figure 4.1). Thus, the dimensionality of discrete decoder output is 5 x T and dimensionality of continuous decoder output is 1 x T, where T is the number of samples. For our 4 sec trials, $T = 49$.

ECoG data: Noise common to all channels was removed by common average referencing. Signals from each channel were band pass filtered with windowed sinc type I linear phase filters of 4 Hz bandwidth between 4-140 Hz (i.e. 4-8, 8-12 .... 136-140 Hz). After excluding bands that include line noise and its harmonic frequencies (i.e. 56-60, 60-64, 64-68, 116-120, 120-124, 124-128), 28 frequency bands were used. Filtered signals were z-scored and Hilbert amplitude of the z-scoed signal is smoothened with a 80 ms moving average window and downsampled to 12.5 Hz. For each trial, ECoG data is restructured into $T \times C \times 28 \times 1$ format where T is the number of time steps, C is the number of electrodes in sensory-motor cortex (SMC), 28 is the number of frequency bands.

4.2.3 Network architectures

Subject-specific architecture

The input to the network is ECoG data in $T \times C \times 28 \times 1$ format. The decoding architecture consists of two time distributed 1 D convolutional layers that find the spatial and frequency filters at each time step in the ECoG input. Then a Long Short Term Memory (LSTM) layer, which is a variant of a recurrent

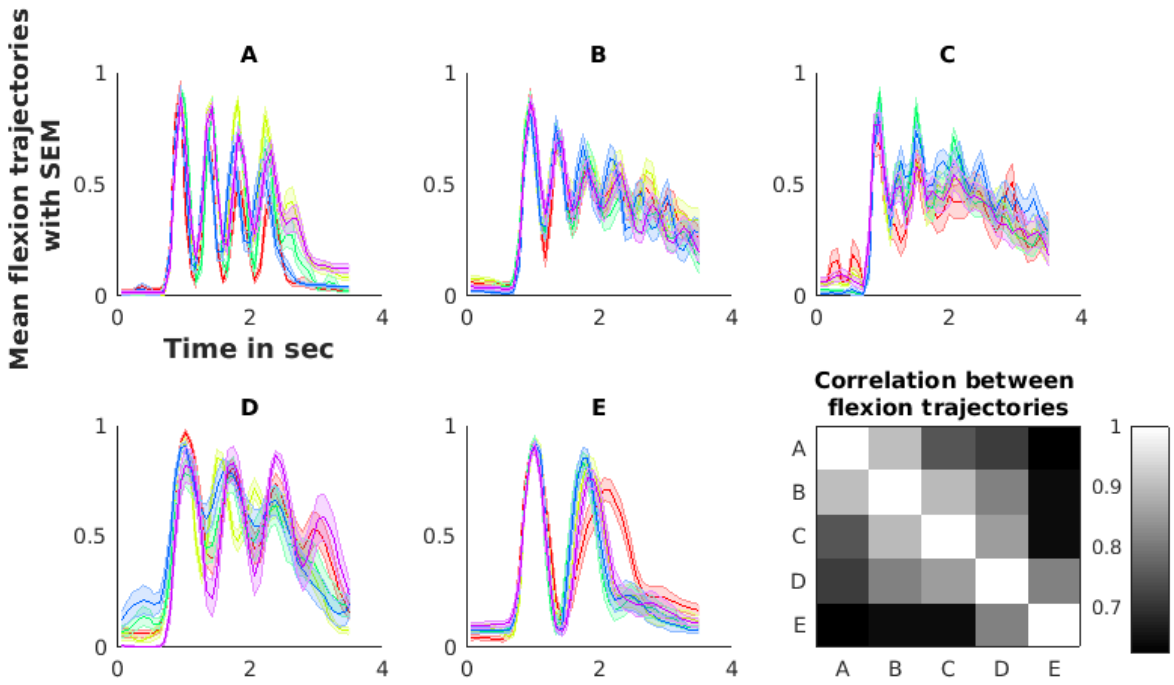

Figure 4.2: Average Flexion traces (with standard errors) of each finger for subjects 'A'-'E'. The bottom right subplot shows the mean corrrelation, averaged across fingers, between subjects.

neural network, is used to learn the temporal dynamics. A time-distributed Dense layer is used to map the output of LSTM to continuous output using an ReLU activation. Another time-distributed Dense layer is used to get discrete output using a softmax activation. A Parameteric ReLU activation layer, dropout layer with a dropout rate of 0.2 (i.e 1 in 5 input nodes will be randomly excluded while training) and a batch normalization layer are applied at the end of convolutional and LSTM layers.The schematic of network architecture is shown in figure 4.3 and is detailed in Table 4.1. A single flexion trajectory is learnt for each trial and the discrete decoder predicts which one of five fingers is moving at each time step of trial. The effective loss function is the sum of mean average error loss for regression and categorical cross entropy loss for discrete decoding. For all the experiments, hidden node size is 10 in each layer of the network and batch size is fixed at 10. Both the regression and discrete outputs are trained in parallel.

4.2.4 Cross-subject architecture

The assumption behind training a cross subject model is that by learning common latent spaces across subjects, we can train a common decoding model that could map the neural activity to behaviors. We propose learning this common latent space in an end to end way using a similar architecture described in section 4.2.3. We call the stacked 1D convolutional layers that convert ECoG activity to input of LSTM layer "Aligner layer". This aligner layer feeds into a common LSTM layer and decoder layers that are fixed across subjects. We train model shown in figure 4.4 in parallel for all subjects. This enables learning aligner layer weights such that the ECoG activity for different subjects maps to a common latent space and the LSTM and decoder layers learn a mapping to the behavioral output. This mapping will thus be generalizable across subjects. For a new test subject, we fix the LSTM and decoder layers learnt using cross-subject data and learn the aligner layer weights to map the subject specific data to previously learnt aligner space.

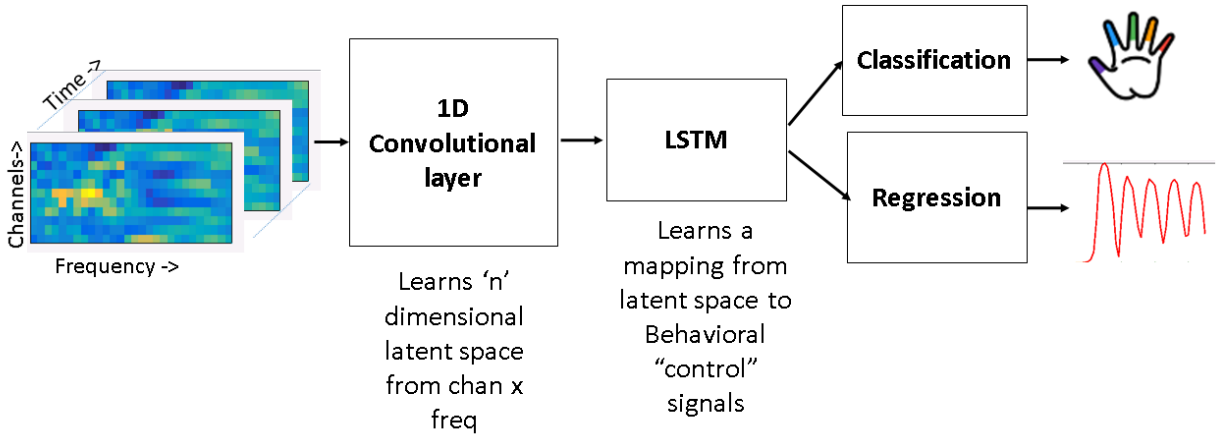

Figure 4.3: Schematic of neural network architecture.

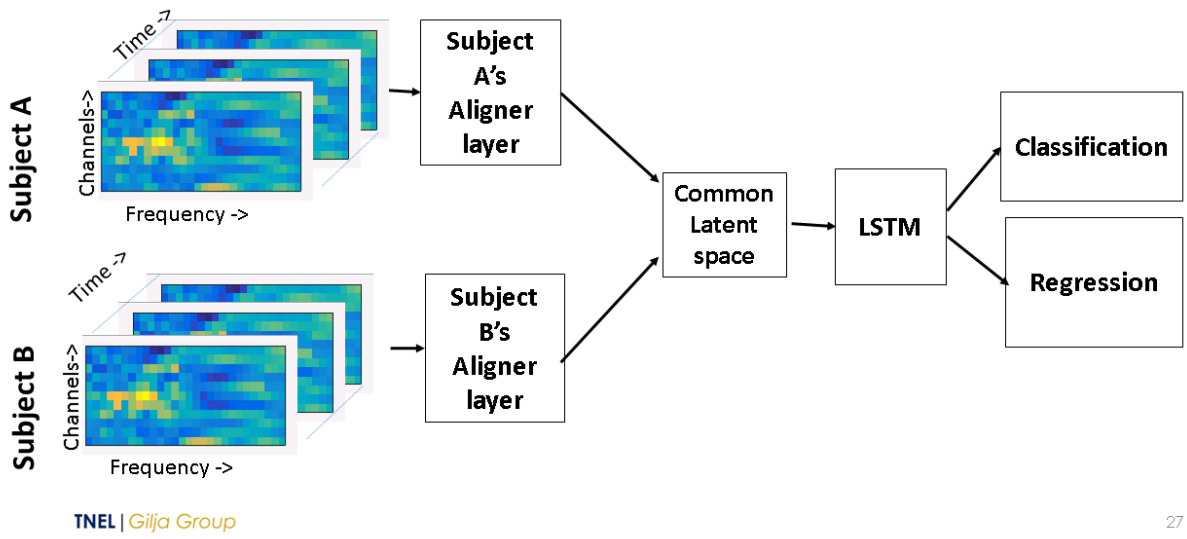

Figure 4.4: Schematic of neural network architecture.

4.3 Results

4.3.1 Learning common latent spaces

The central idea behind training cross-subject models is that we can learn an alignment between neural latent spaces across subjects. We achieve that by fixing the decoder, i.e. the LSTM + dense layers, that maps from latent space to behavioral output and letting the convolutional layer weights be trained in such a way that the input to decoders are aligned across subjects. Figure 4.5 shows the first three principal components of hidden node activations of the aligner layer of subject B learnt to align with the latent space of model learnt usng data from subject A. This shows that the network architecture allows us to learn common latent spaces in a supervised way.

Table 4.1: Network architecture for ECoG decoding

Layer	Input Shape	Output Shape	Notes on layer
Input Layer	(49,N,28,1)	(49,N,28,1)	4 sec ECoG signal
Time distributed 1D Conv layer	(49,N,28,1)	(49,N,1,F1)	$F1$ is number of filters to model frequency patterns
Time distributed 1D Conv layer	(49,N,1,F1)	(49,1,1,F2)	$F2$ is number of filters to model spatial patterns
PReLU activation	(49,1,1,F2)	(49,1,1,F2)	Parametric ReLU activation
Dropout	(49,1,1,F2)	(49,1,1,F2)	dropout rate of 0.2
Batch Normalization	(49,1,1,F2)	(49,1,1,F2)	batch normalization layer for addressing internal covariate shift
Time Distributed Flatten layer	(49,1,1,F2)	(49,F2)	Flatten the output for input to LSTM
LSTM layer	(49,F2)	(49,F3)	$F3$ is the number of hidden nodes in LSTM layer
Dropout	(49,F3)	(49,F3)	
Batch Normalization	(49,F3)	(49,F3)	
Time distributed Dense layer (ReLU activation)	(49,1)	(49,1)	Predict continuous flexion trajectory at each time step
Time distributed Dense layer (sigmoid activation)	(49,5)	(49,5)	Predict moving finger at each time step

4.3.2 Decoding results

Figure 4.6 shows the discrete classification accuracy, mean average error and correlation coefficient between true and predicted finger flexion trajectories. For calculating the classification accuracy from the $5 \times T$ dimensional output of discrete decoder we calculate the mean prediction probabilities of each finger during the flexion period and assign the finger that has maximum value as predicted finger. Cross-subject models perform on par with subject-specific models. Figure 4.7 shows qualitatively the actual and predicted flexion trajectories. By fixing decoders and only learning the aligner layer weights, we can reduce the number of training epochs required to train the model for a new test subject as shown in Figure 4.8. The figure also shows that cross-subject models are less prone to overfitting compared to subject-specific models. Figure 4.9 shows that by leveraging information across subjects, the LSTM layer is better able to model dynamics to seperate the fingers.

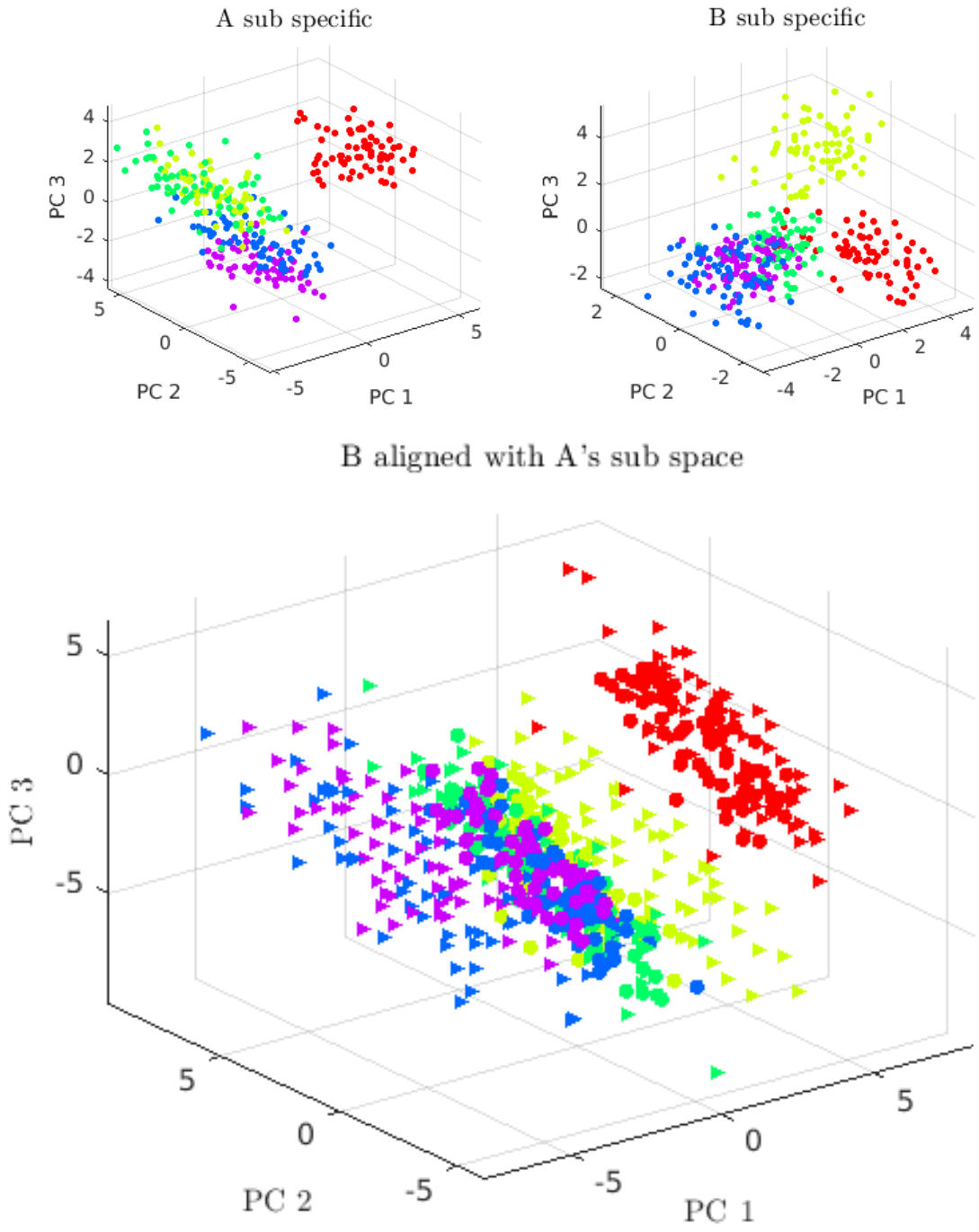

Figure 4.5: Common latent spaces learnt using the neural network at the end of aligner network. Latent spaces of subject specific models (top). Subject B's latent space is aligned with subject A's space by fixing the LSTM and dense layers with subject A's model weights (bottom). Colors correspond to different fingers and shapes correspond to subjects

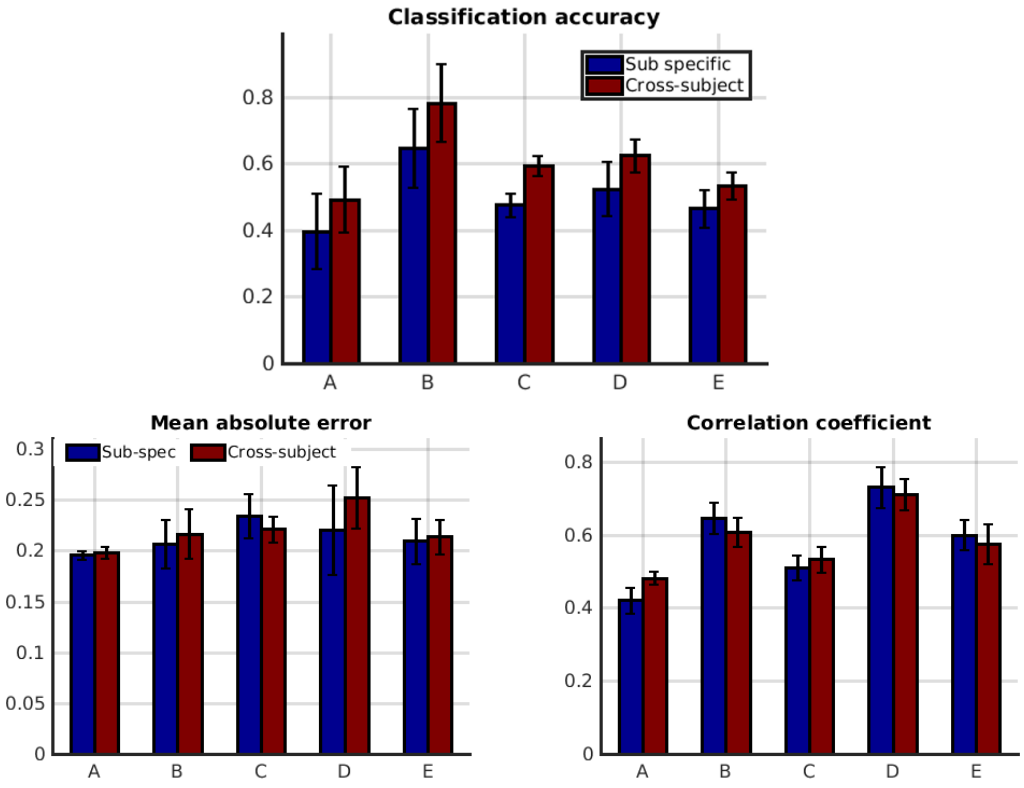

Figure 4.6: Continuous and discrete decoding results comparing subject specific and cross-subject models.

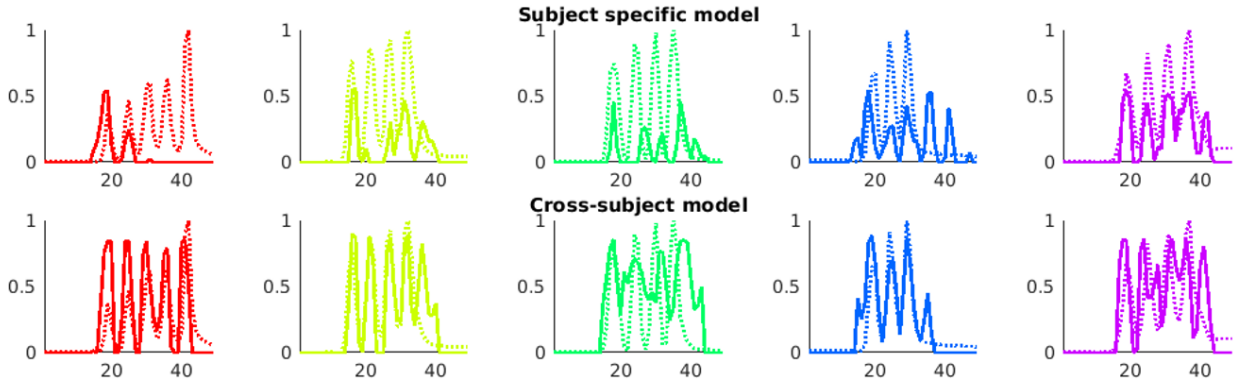

Figure 4.7: Sample trials for subject A showing predicted and true flexions for subject specific and cross-subject model for each finger (columns).

4.4 Discussion

Motivation to explore cross-subject learning for ECoG based BMIs is three fold. Firstly, cross-subject learning allows us to leverage data from multiple subjects to decode for new subject. Secondly, this could help reduce the time required to calibrate BMI decoder for new session or user. Thirdly, this would enable us to use the dynamics learnt from existing data corpus to build prosthesis that can perform tasks in more degrees of freedom than the user is capable of. In this work, we showed that it is possible to learnt common models using cross-subject data and these models perform on par with subject-specific models.

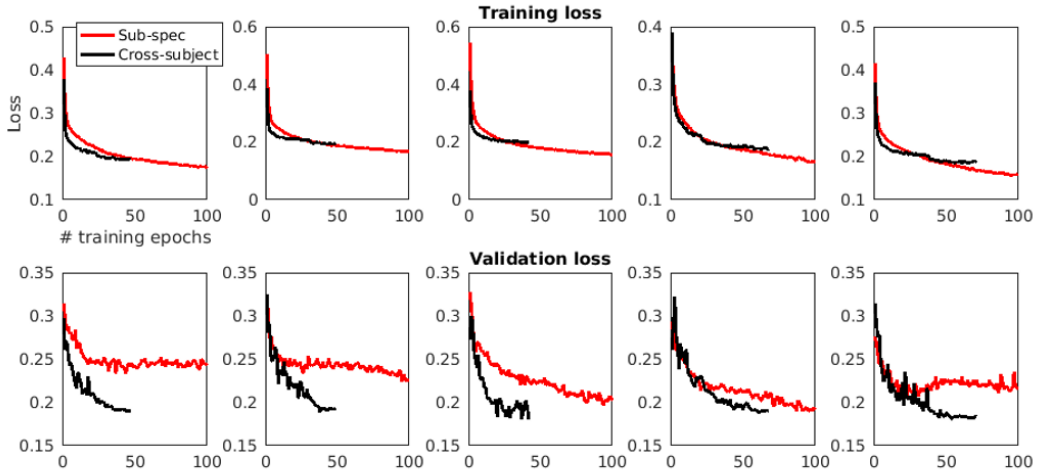

Figure 4.8: Comparing training and validation losses of subject specific and cross subject models

Chapter 4, in part, is currently being prepared for submission for publication of the material. "Cross-subject learning for decoding electrocorticographic signals". Pailla, Tejaswy, Venkatesh Elango, Aashish Patel, Vikash Gilja. The dissertation author is the primary investigator and author of this paper.

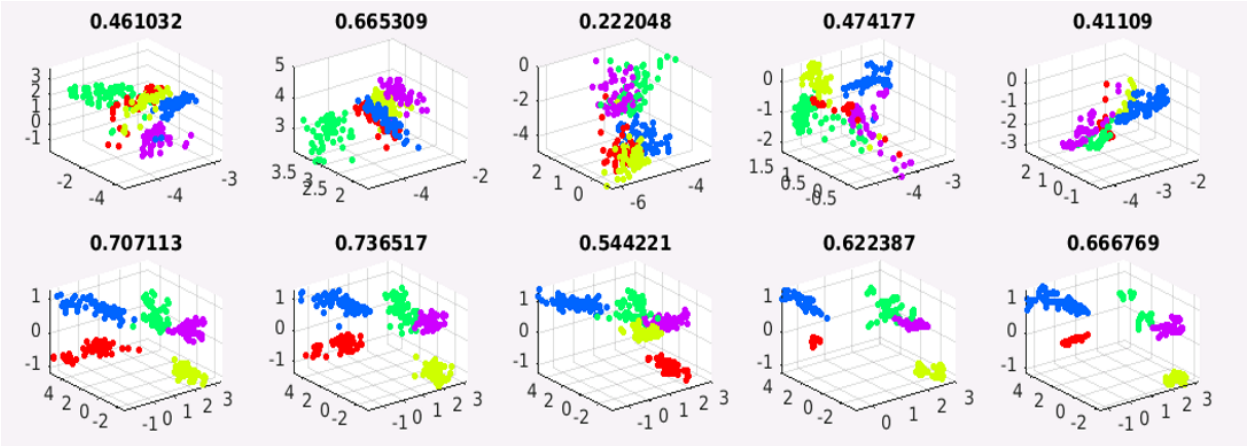

Figure 4.9: Comparing cluster separability of hidden node activation of LSTM output for subject E. The five columns correspond to five cross validation folds. Colors correspond to different fingers and the title for each subplot is the Silhouette cluster separability index. We can see that the hidden node activations of cross-subject models are more separable than subject-specific models

References

[1] Edwin M Maynard, Craig T Nordhausen, and Richard A Normann. The utah intracortical electrode array: a recording structure for potential brain-computer interfaces. *Electroencephalography and clinical neurophysiology*, 102(3):228–239, 1997.

[2] MD Serruya, NG Hatsopoulos, L Paninski, MR Fellows, and JP Donoghue. Instant neural control of a movement signal. hands-free operation of a cursor can be achieved by a few neurons in the motor cortex. *Nature*, 416:141–142, 2002.

[3] Leigh R Hochberg, Mijail D Serruya, Gerhard M Friehs, Jon A Mukand, Maryam Saleh, Abraham H Caplan, Almut Branner, David Chen, Richard D Penn, and John P Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099):164, 2006.

[4] Leigh R Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y Masse, John D Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S Cash, Patrick Van Der Smagt, et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372, 2012.

[5] Eric C Leuthardt, Kai J Miller, Gerwin Schalk, Rajesh PN Rao, and Jeffrey G Ojemann. Electrocorticography-based brain computer interface-the seattle experience. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):194–198, 2006.

[6] G Schalk, J Kubanek, KJ Miller, NR Anderson, EC Leuthardt, JG Ojemann, Dave Limbrick, D Moran, LA Gerhardt, and JR Wolpaw. Decoding two-dimensional movement trajectories using electrocorticographic signals in humans. *Journal of neural engineering*, 4(3):264, 2007.

[7] Gerwin Schalk. Can electrocorticography (ecog) support robust and powerful brain-computer interfaces? *Frontiers in neuroengineering*, 3:9, 2010.

[8] Eric C Leuthardt, Gerwin Schalk, Jonathan R Wolpaw, Jeffrey G Ojemann, and Daniel W Moran. A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, 1(2):63, 2004.

[9] Guy Hotson, David P McMullen, Matthew S Fifer, Matthew S Johannes, Kapil D Katyal, Matthew P Para, Robert Armiger, William S Anderson, Nitish V Thakor, Brock A Wester, et al. Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject. *Journal of neural engineering*, 13(2):026017, 2016.

[10] Wei Wang, Jennifer L Collinger, Alan D Degenhart, Elizabeth C Tyler-Kabara, Andrew B Schwartz, Daniel W Moran, Douglas J Weber, Brian Wodlinger, Ramana K Vinjamuri, Robin C Ashmore, et al. An electrocorticographic brain interface in an individual with tetraplegia. *PloS one*, 8(2):e55344, 2013.

[11] Alan D Degenhart, Shivayogi V Hiremath, Ying Yang, Stephen Foldes, Jennifer L Collinger, Michael Boninger, Elizabeth C Tyler-Kabara, and Wei Wang. Remapping cortical modulation for electrocorticographic brain–computer interfaces: a somatotopy-based approach in individuals with upper-limb paralysis. *Journal of neural engineering*, 15(2):026021, 2018.

[12] Takufumi Yanagisawa, Masayuki Hirata, Youichi Saitoh, Haruhiko Kishima, Kojiro Matsushita, Tetsu Goto, Ryohei Fukuma, Hiroshi Yokoi, Yukiyasu Kamitani, and Toshiki Yoshimine. Electrocorticographic control of a prosthetic arm in paralyzed patients. *Annals of neurology*, 71(3):353–361, 2012.

[13] Zenas C Chao, Yasuo Nagasaka, and Naotaka Fujii. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkey. *Frontiers in neuroengineering*, 3:3, 2010.

[14] Jeremiah D Wander, Timothy Blakely, Kai J Miller, Kurt E Weaver, Lise A Johnson, Jared D Olson, Eberhard E Fetz, Rajesh PN Rao, and Jeffrey G Ojemann. Distributed cortical adaptation during learning of a brain–computer interface task. *Proceedings of the National Academy of Sciences*, 110(26):10818–10823, 2013.

[15] James C Barrese, Naveen Rao, Kaivon Paroo, Corey Triebwasser, Carlos Vargas-Irwin, Lachlan Franquemont, and John P Donoghue. Failure mode analysis of silicon-based intracortical microelectrode arrays in non-human primates. *Journal of neural engineering*, 10(6):066014, 2013.

[16] F Schettini, F Aloise, P Aricò, S Salinari, D Mattia, and F Cincotti. Self-calibration algorithm in an asynchronous p300-based brain–computer interface. *Journal of neural engineering*, 11(3):035004, 2014.

[17] Ali Farshchian, Juan Alvaro Gallego, Joseph Paul Cohen, Yoshua Bengio, Lee E. Miller, and Sara A. Solla. Adversarial domain adaptation for stable brain-machine interfaces. *ArXiv*, abs/1810.00045, 2018.

[18] Eva L Dyer, Mohammad Gheshlaghi Azar, Matthew G Perich, Hugo L Fernandes, Stephanie Naufel, Lee E Miller, and Konrad P Körding. A cryptography-based approach for movement decoding. *Nature biomedical engineering*, 1(12):967, 2017.

[19] Minmin Cheng, Zuhong Lu, and Haixian Wang. Regularized common spatial patterns with subject-to-subject transfer of eeg signals. *Cognitive neurodynamics*, 11(2):173–181, 2017.

[20] Mark Wronkiewicz, Eric Larson, and Adrian KC Lee. Leveraging anatomical information to improve transfer learning in brain–computer interfaces. *Journal of neural engineering*, 12(4):046027, 2015.

[21] Vinay Jayaram, Morteza Alamgir, Yasemin Altun, Bernhard Scholkopf, and Moritz Grosse-Wentrup. Transfer learning in brain-computer interfaces. *IEEE Computational Intelligence Magazine*, 11(1):20–31, 2016.

[22] Siavash Sakhavi and Cuntai Guan. Convolutional neural network-based transfer learning and knowledge distillation using multi-subject data in motor imagery bci. In *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 588–591. IEEE, 2017.

[23] Fatemeh Fahimi, Zhuo Zhang, Wooi Boon Goh, Tih-Shih Lee, Kai Keng Ang, and Cuntai Guan. Inter-subject transfer learning with end-to-end deep convolutional neural network for eeg-based bci. *Journal of neural engineering*, 2018.

[24] Kai J Miller, Dora Hermes, Christopher J Honey, Adam O Hebb, Nick F Ramsey, Robert T Knight, Jeffrey G Ojemann, and Eberhard E Fetz. Human motor cortical activity is selectively phase-entrained on underlying rhythms. *PLoS Comput Biol*, 8(9):e1002655, 2012.

[25] KJ Miller and JG Ojemann. A library of human electrocorticographic data and analyses. *Program No. 469.08. 2016 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2016. Online*.

[26] Tejaswy Pailla, Kai Joshua Miller, and Vikash Gilja. Autoencoders for learning template spectrograms in electrocorticographic signals. *Journal of neural engineering*, 2018.

Chapter 5

Future work

5.1 Analyzing large-scale ECoG data

Until recently, ECoG based research studies have largely relied on datasets collected from subjects who volunteered to participate in experiments while undergoing treatment in epilepsy monitoring units (EMUs). These datasets are typically of the order of tens of minutes where the subjects perform controlled experimental tasks, resulting in well-labelled, relatively clean datasets. The experiment and subject choices are usually done based on the ECoG coverage and the subject's ability or willingness to do tasks. Such studies have largely contributed to the understanding of specific cortical areas and brain-machine interface studies.

However, recently researchers are pursuing the opportunity to collect ECoG data for extended periods of times, i.e. usually in the order of days, while the subjects are in EMUs. Behavioral data, i.e. the subject's physical activity or movements are recorded using non-invasive or minimally invasive sensors like Kinect that capture audio, video [1, 2]. The volume of these datasets are large compared to the few minutes of datasets used in research studies til date. Furthermore, ECoG data collected under naturalistic conditions is noisy and coarsely labelled making it difficult to study event related potentials. The traditional machine learning methods used for smaller datasets do not scale well with these large datasets. Therefore, there is

now a need to develop scalable and unsupervised toolkit to analyze these novel datasets.

5.2 Proof-of-concept results

I use the data presented in [3], where four contexts; Rest, Dialogue, Electronics, TV; have been coarsely labelled for three subjects: NY394, RCH1 and RCH3. Figure 5.1 shows the timeline of contexts for three subjects. While subject NY394 has only ECOG electrodes, RCH1 and RCH3 have both ECoG and sEEG coverage. ECoG channels were common average referenced and sEEG channels were bipolar referenced. These data were collected over 3-4 days. The data were coarsely labelled in 5 to 15 min segments.

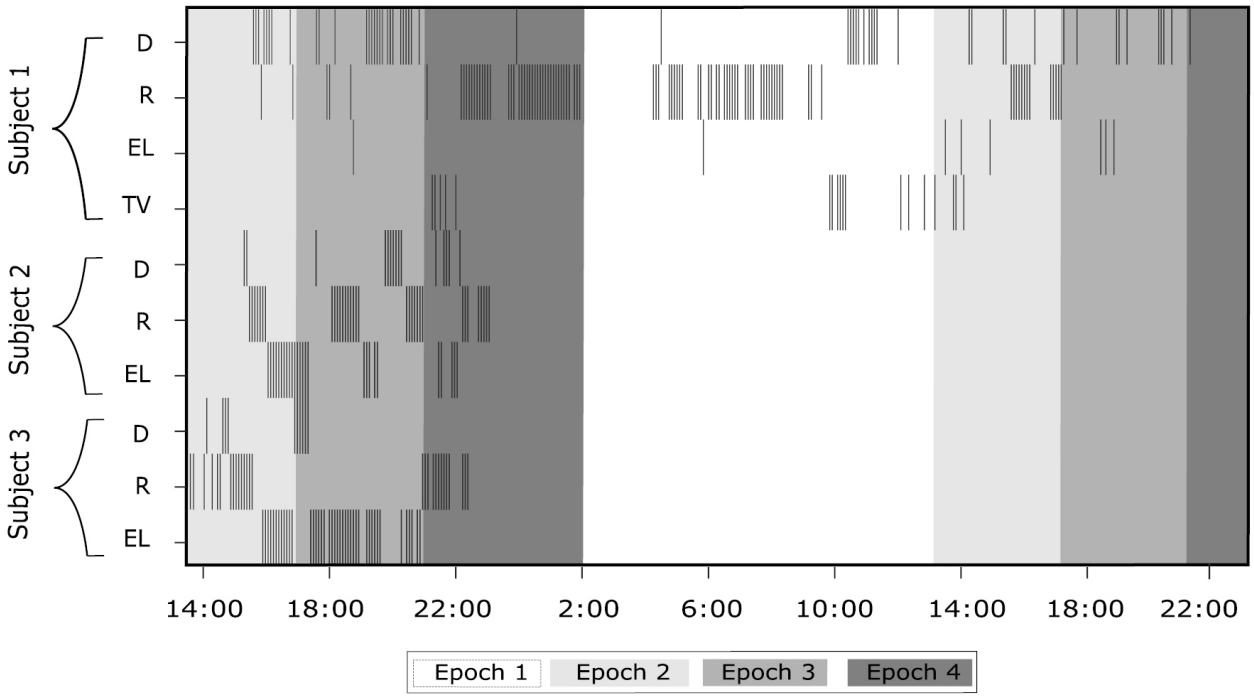

Figure 5.1: Raster plot showing the locations of each context in time: Rest (R), Dialogue (D), Electronics (EL), Television (TV). Source [3]

5.2.1 Unsupervised methods to cluster ECoG data

Variational autoencoders (VAE) are generational models that learn latent representations of data and allow sampling new samples from this latent distributions [4]. A two layer VAE with 100 hidden nodes in first layer and 2 hidden nodes in second layer was used to find the latent representation of data. The resulting latent space was clustered using a gaussian mixture model and compared with true labels. The existence of mutual information between the true and clustered labels as shown in 5.2 demonstrate that the VAE could extract spatio-temporal patterns in the data that could extract context relevant latent representations from the neural data. This kind of unsupervised clustering methods can further be explored to automate labelling in large scale datasets.

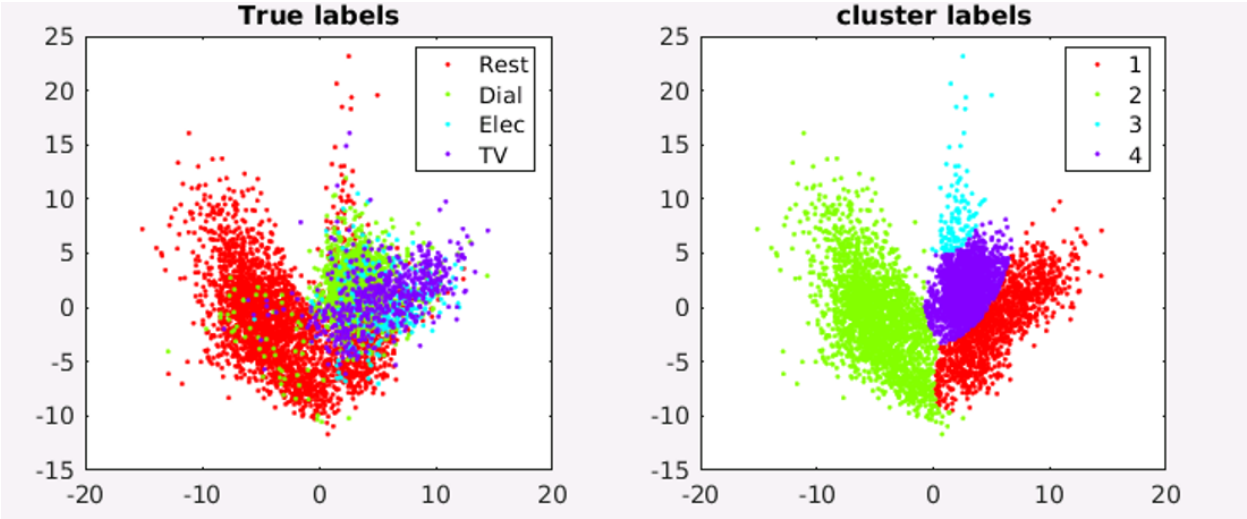

Figure 5.2: Clustering contexts using variational autoencoders for NY394.

5.3 Future Directions

A major challenge in analyzing large-scale data is defining the scope of the questions that can be answered with these multi-modal datasets. Given the variation in grid coverage, exact location of electrodes on the cortical surface and lack of event related timings, it might not be possible to study functional orga-

nization of particular cortical areas pertaining to nuanced behaviors. However, these datasets can help us answer more broader questions such as connectivity of cortical regions, finding spatial-temporal patterns representative of abstract brain states such as when a subject is at rest or engaging in dialogue. While decoding coarse behavioral states (e.g. rest vs dialogue) from naturalistic conditions can find applications in BMIs [3], these datasets could be much more valuable to neuroscientists seeking to understand connectivity patterns of brain in the wild.

References

[1] Nancy XR Wang, Ali Farhadi, Rajesh PN Rao, and Bingni W Brunton. Ajile movement prediction: Multimodal deep learning for natural human neural recordings and video. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[2] Paolo Gabriel, Werner K Doyle, Orrin Devinsky, Daniel Friedman, Thomas Thesen, and Vikash Gilja. Neural correlates to automatic behavior estimations from rgb-d video in epilepsy unit. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3402–3405. IEEE, 2016.

[3] Abdulwahab Alasfour, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orin Devinsky, et al. Coarse behavioral context decoding. *Journal of neural engineering*, 16(1):016021, 2019.

[4] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.