Briefings in Bioinformatics, 22(2), 2021, 1577–1591

doi: 10.1093/bib/bbaa355
Advance Access Publication Date: 29 December 2020
Method Review

Deep learning approaches for neural decoding across architectures and recording modalities

Jesse A. Livezey and Joshua I. Glaser

Corresponding authors: Jesse Livezey, Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States; Redwood Center for Theoretical Neuroscience, University of California, Berkeley, Berkeley, California, United States. E-mail: jlivezey@lbl.gov; Joshua Glaser, Department of Statistics, Columbia University, New York, United States; Zuckerman Mind Brain Behavior Institute, Columbia University, New York, United States; Center for Theoretical Neuroscience, Columbia University, New York, United States; E-mail: j.glaser@columbia.edu

Abstract

Decoding behavior, perception or cognitive state directly from neural signals is critical for brain–computer interface research and an important tool for systems neuroscience. In the last decade, deep learning has become the state-of-the-art method in many machine learning tasks ranging from speech recognition to image segmentation. The success of deep networks in other domains has led to a new wave of applications in neuroscience. In this article, we review deep learning approaches to neural decoding. We describe the architectures used for extracting useful features from neural recording modalities ranging from spikes to functional magnetic resonance imaging. Furthermore, we explore how deep learning has been leveraged to predict common outputs including movement, speech and vision, with a focus on how pretrained deep networks can be incorporated as priors for complex decoding targets like acoustic speech or images. Deep learning has been shown to be a useful tool for improving the accuracy and flexibility of neural decoding across a wide range of tasks, and we point out areas for future scientific development.

Introduction

Using signals from the brain to make predictions about behavior, perception or cognitive state, i.e. 'neural decoding', is becoming increasingly important within neuroscience and engineering. One common goal of neural decoding is to create brain computer interfaces, where neural signals are used to control an output in real time [1, 2]. This could allow patients with neurological or motor diseases or injuries to, for example, control a robotic arm or cursor on a screen, or produce speech through a synthesizer. Another common goal of neural decoding is to gain a better scientific understanding of the link between neural activity and the outside world. To provide insight, decoding accuracy can be compared across brain regions, cell types, different types of subjects (e.g. with different diseases or genetics) and different experimental conditions [3–11]. In addition, the representations learned by neural decoders can be probed to better understand the structure of neural computation [12–16]. These uses of neural decoding span many different neural recording modalities and span a wide range of behavioral outputs (Figure 1A).

Within the last decade, many researchers have begun to successfully use deep learning approaches for neural decoding. A decoder can be thought of as a function approximator, doing either regression or classification depending on whether the output is a continuous or categorical variable. Given the great successes of deep learning at learning complex functions across many domains [17–26], it is unsurprising that deep learning has become a popular approach in neuroscience. Here, we review the many uses of deep learning for neural decoding. We emphasize how different deep learning architectures can induce biases that can be beneficial when decoding from different neural recording modalities and when decoding different behavioral outputs. We aim to provide a review that is both useful to deep learning researchers looking to understand current neural decoding problems and to neuroscience researchers looking to understand the state-of-the-art in neural decoding.

Jesse A. Livezey is a postdoctoral research scientist in the Neural Systems and Data Science Laboratory at the Lawrence Berkeley National Laboratory. He obtained his PhD in Physics from the University of California, Berkeley. His research interests include applications of machine learning and information theory to neuroscience datasets.
Joshua I. Glaser is a postdoctoral research scientist in the Center for Theoretical Neuroscience and Department of Statistics at Columbia University. He obtained his PhD in Neuroscience from Northwestern University. His research interests include neuroscience, machine learning and statistics.
Submitted: 18 May 2020; **Received (in revised form):** 31 October 2020