Journal of Neuroscience Methods 294 (2018) 72–80

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Detection of relationships among multi-modal brain imaging meta-features via information flow

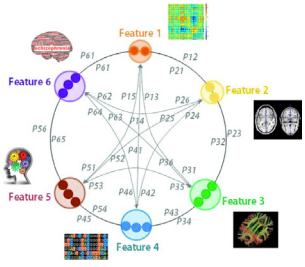
Robyn L. Miller[a,*], Victor M. Vergara[a], Vince D. Calhoun[a,b]

[a] The Mind Research Network, 1101 Yale Blvd. NE, Albuquerque, NM, 87131, USA
[b] Department of Electrical and Computer Engineering, 498 Terrace St. NE, Albuquerque, NM, 87106, USA

HIGHLIGHTS

- Proposed method yields a high-level summary of all relationships between data features in large multimodal studies.
- Jointly accommodates data from very different modalities: brain imaging, behavioral, biological/genetic etc.
- Data features can have any dimension.
- Relationships are not assumed to be linear.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 7 November 2017
Accepted 9 November 2017
Available online 14 November 2017

Keywords:
Big data
Biomedical imaging
Biomedical signal processing
Data fusion
Multimodal brain imaging

ABSTRACT

Background: Neuroscientists and clinical researchers are awash in data from an ever-growing number of imaging and other bio-behavioral modalities. This flow of brain imaging data, taken under resting and various task conditions, combines with available cognitive measures, behavioral information, genetic data plus other potentially salient biomedical and environmental information to create a rich but diffuse data landscape. The conditions being studied with brain imaging data are often extremely complex and it is common for researchers to employ more than one imaging, behavioral or biological data modality (e.g., genetics) in their investigations. While the field has advanced significantly in its approach to multimodal data, the vast majority of studies still ignore joint information among two or more features or modalities.
New method: We propose an intuitive framework based on conditional probabilities for understanding information exchange between features in what we are calling a feature meta-space; that is, a space consisting of many individual featurae spaces. Features can have any dimension and can be drawn from any data source or modality. No a priori assumptions are made about the functional form (e.g., linear, polynomial, exponential) of captured inter-feature relationships.
Results: We demonstrate the framework's ability to identify relationships between disparate features of varying dimensionality by applying it to a large multi-site, multi-modal clinical dataset, balance between schizophrenia patients and controls. In our application it exposes both expected (previously observed) relationships, and novel relationships rarely considered investigated by clinical researchers.
Comparison with existing method(s): To the best of our knowledge there is not presently a comparably efficient way to capture relationships of indeterminate functional form between features of arbitrary dimension and type. We are introducing this method as an initial foray into a space that remains relatively underpopulated.

* Corresponding author.
E-mail address: rmiller@mrn.org (R.L. Miller).

https://doi.org/10.1016/j.jneumeth.2017.11.006

Conclusions: The framework we propose is powerful, intuitive and very efficiently provides a high-level overview of a massive data space. In our application it exposes both expected relationships and relationships very rarely considered worth investigating by clinical researchers.

1. Introduction

Working in meta-spaces, i.e. evaluating relationships among summaries of individual feature spaces, presents many challenges, the most obvious of which is the sheer size of the space and the incommensurable content and dimensionality of constituent features. Productively integrating large numbers of multimodal features offers the promise of identifying intricate signatures of very complex symptomatically overlapping conditions, all within a multimodal information landscape (Bießmann et al., 2011; Calhoun and Adali, 2009; Friston, 2009; Lahat et al., 2015; Schultz et al., 2012; Sui et al., 2012, 2014, 2015, 2013).

Multimodal-multivariate fusion techniques promise to be a better approach to analyze biological data than previous univariate-single modal approaches (Calhoun and Sui, 2016). One of the main advantages is the ability of bringing together complementary information from several modalities to improve their analysis (Calhoun and Adali, 2009). However, such approaches are challenging because of their higher complexity and dimensionality (Sui et al., 2012). In addition, there is no numerical method to date that quantifies to which degree information in one modality is related to that found in a complementary modality. Such quantification can provide important evidence to support the use of multimodal techniques.

We propose a model of directional information flow between feature spaces of arbitrary dimension, applicable to any number of features from any modality (Fig. 1). Advantages include computational efficiency, tolerance of missing data, flexible incorporation of features at multiple dimensional scales (e.g., scalar coupling strength between a single pair of brain networks vs. coupling strengths between networks in two broad functional domains vs. coupling strengths between all brain networks) and smooth incorporation of emerging imaging modalities. Relationships captured between features and modalities are intrinsically multivariate and need not be linear (Fig. 2(B)). Captured relationships are helpful in providing an information map providing numerical evidence for considering links among different data modalities.

2. Methods

We induce a weighted directed graph between features of arbitrary dimension based on the degree to which the distributional geometry of target features is conditioned upon (informed by) the distributional geometry of source features. Each feature space is divided into clusters, taken to be the "open sets" of these features as probability spaces. The result is a mapping of probabilities among features (Fig. 1).

2.1. Conditional probabilities between features

For each ordered pair F_i, $F_j$of features $i, j \in \{1, 2, \ldots, k\}$ from the set of k features being analyzed, we compute

$$P_{i,j} = \left(p_{r,s}^{(i,j)}\right)_{r,s=1}^{n_i,n_j},$$

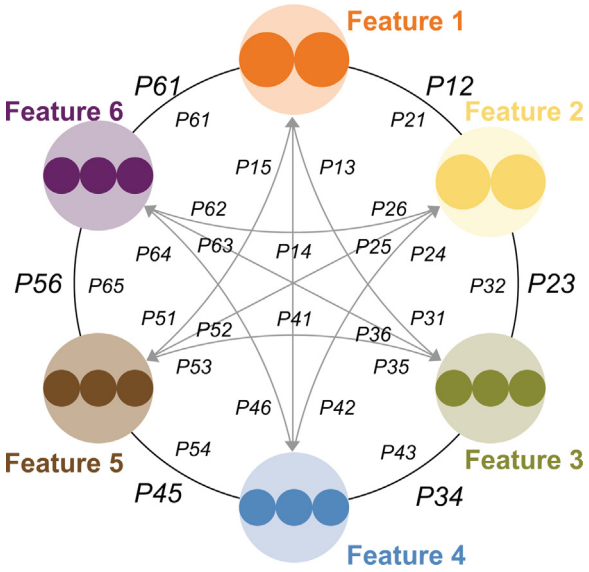


Fig. 1. Illustration of directional mapping via conditional probability matrices between clustered data features of arbitrary type and dimension.

the (population-wide) probabilities

$$p_{r,s}^{(i,j)} = p\left(C_s^{(j)} | C_r^{(i)}\right)$$

of being in the s^{th}, $s \in \{1, .., n_j\}$ cluster, $C_s^{(j)}$, of F_j *conditional on* being in the r^{th}, $r \in \{1, .., n_i\}$ cluster of F_i ($i \neq j$). These conditional probability matrices, referred to as *transition probabilities*, row-sum to one and act like Markov transition matrices between feature pair cluster structures.

2.2. Information flow between features

We consider a source feature F_i to be *informative* about a target feature F_j if the probabilities of occupying clusters in the target depend heavily on which cluster of the source is being conditioned upon.

There are two criteria we require for the metric $\mathcal{J}_{i,j}$ of information flow from F_i to F_j: it should increase in (i) how dissimilar the full distributions on target

clusters from source clusters and also increase in (ii) how well clusters of the target separate clusters of the source. The upper boundary of distributional dissimilarity will show good separation. To ensure this is also true away from this extreme case we introduce two informational measures, each tuned to one of the objectives: $\mathcal{D}_{i,j} \in [0, 1]$ for the degree of broad distributional dissimilarity and $\mathcal{S}_{i,j} \in [0, 1]$ for the degree to which source clusters map preferentially to some specific target cluster. The overall information metric $\mathcal{J}_{i,j} \in [0, 1]$ is the average of $\mathcal{D}_{i,j}$ and $\mathcal{S}_{i,j}$. The L^2 distances between rows of $P_{i,j}$ are a reasonable indicator of distributional dissimilarity. Each of the pairwise L^2 distances is bounded in $[0, \sqrt{2}]$. We average