Spectral Graph Convolutions for Population-based Disease Prediction

Sarah Parisot⋆, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee,
Ricardo Guerrerro Moreno, Ben Glocker, and Daniel Rueckert

Biomedical Image Analysis Group, Imperial College London, UK

Abstract. Exploiting the wealth of imaging and non-imaging information for disease prediction tasks requires models capable of representing, at the same time, individual features as well as data associations between subjects from potentially large populations. Graphs provide a natural framework for such tasks, yet previous graph-based approaches focus on pairwise similarities without modelling the subjects' individual characteristics and features. On the other hand, relying solely on subject-specific imaging feature vectors fails to model the interaction and similarity between subjects, which can reduce performance. In this paper, we introduce the novel concept of Graph Convolutional Networks (GCN) for brain analysis in populations, combining imaging and non-imaging data. We represent populations as a sparse graph where its vertices are associated with image-based feature vectors and the edges encode phenotypic information. This structure was used to train a GCN model on partially labelled graphs, aiming to infer the classes of unlabelled nodes from the node features and pairwise associations between subjects. We demonstrate the potential of the method on the challenging ADNI and ABIDE databases, as a proof of concept of the benefit from integrating contextual information in classification tasks. This has a clear impact on the quality of the predictions, leading to 69.5% accuracy for ABIDE (outperforming the current state of the art of 66.8%) and 77% for ADNI for prediction of MCI conversion, significantly outperforming standard linear classifiers where only individual features are considered.

1 Introduction

Recent years have seen an increasing volume of medical image data being collected and stored. Large scale collaborative initiatives are acquiring and sharing hundreds of terabytes of imaging, genetic and behavioural data. With this novel wealth of imaging and non-imaging data, there is a need for models capable of representing potentially large populations and exploiting all types of information. Graphs provide a natural way of representing populations and their similarities. In such setting, each subject acquisition is represented by a node and pairwise similarities are modelled via weighted edges connecting the nodes.

Such models provide powerful tools for population analysis and integration of non-imaging data such as manifold learning [2,15] or clustering algorithms [11]. Nonetheless, all the available information is encoded via pairwise similarities, without modelling the subjects' individual characteristics and features. On the other hand, relying solely on imaging feature vectors, e.g. to train linear classifiers as in [1], fails to model the interaction and similarity between subjects. This can make generalisation more difficult and reduce performance, in particular when the data is acquired using different imaging protocols. Convolutional Neural Networks (CNNs) have found numerous applications on 2D and 3D images, as powerful models that exploit features (e.g. image intensities) and neighbourhood information (e.g. the regular pixel grid) to yield hierarchies of features and solve problems like image segmentation [7] and classification. The task of subject classification in populations (e.g. for diagnosis) can be compared to image segmentation where each pixel is to be classified. In this context, an analogy can be made between an image pixel and its intensity, and a subject and its corresponding feature vectors, while the pairwise population graph equates to the pixel grid, describing the neighbourhood structure for convolutions. However, the application of CNNs on irregular graphs is not straightforward. This requires the definition of local neighbourhood structures and node orderings for convolution and pooling operations [10], which can be challenging for irregular graph structures. Recently, graph CNNs were introduced [4], exploiting the novel concept of signal processing on graphs [13], which uses computational harmonic analysis to analyse signals defined on irregular graph structures. These properties allow convolutions in the graph spatial domain to be dealt as multiplications in the graph spectral domain, extending CNNs to irregular graphs in a principled way. Such graph CNN formulation was successfully used in [8] to perform classification of large citation datasets.

Contributions. In this paper, we introduce the novel concept of Graph Convolutional Networks (GCN) for brain analysis in populations, combining imaging and non-imaging data. Our goal is to leverage the auxiliary information available with the imaging data to integrate similarities between subjects within a graph structure. We represent the population as a graph where each subject is associated with an imaging feature vector and corresponds to a graph vertex. The graph edge weights are derived from phenotypic data, and encode the pairwise similarity between subjects and the local neighbourhood system. This structure is used to train a GCN model on partially labelled graphs, aiming to infer the classes of unlabelled nodes from the node features and pairwise associations between subjects. We demonstrate the potential of the method on two databases, as a proof of concept of the advantages of integrating contextual information in classification tasks. First, we classify subjects from the Autism Brain Imaging Data Exchange (ABIDE) database as healthy or suffering from Autism Spectrum Disorders (ASD). The ABIDE dataset comprises highly heterogeneous functional MRI data acquired at multiple sites. We show how integrating acquisition information allows to outperform the current state of the art on the whole dataset [1] with a global accuracy of 69.5%. Second, using the Alzheimer's Dis-
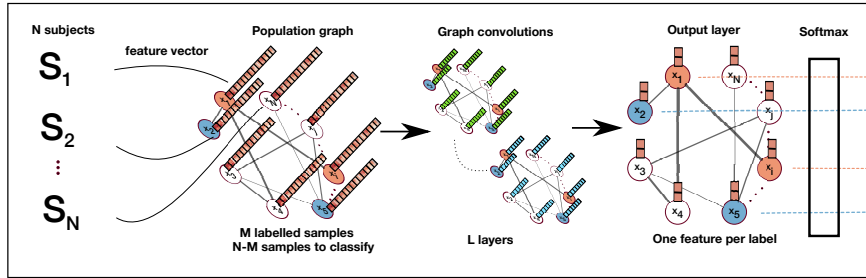
Fig. 1: Overview of the pipeline used for classification of population graphs using Graph Convolutional Networks.

ease Neuroimaging Initiative (ADNI) database, we show how our model allows to seamlessly integrate longitudinal data and provides a significant increase in performance to 77% accuracy for the challenging task of predicting the conversion from Mild Cognitive Impairment (MCI) to Alzheimer's Disease (AD). The code is publicly available at `https://github.com/parisots/population-gcn`.

2   Methods

We consider a database of N acquisitions comprising imaging (e.g. resting-state fMRI or structural MRI) and non-imaging phenotypic data (e.g. age, gender, acquisition site, etc.). Our objective is to assign to each acquisition, corresponding to a subject and time point, a label $l \in \mathscr{L}$ describing the corresponding subject's disease state (e.g. control or diseased). To this end, we represent the population as a sparse graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, W\}$ where W is the adjacency matrix describing the graph's connectivity. Each acquisition S_v is represented by a vertex $v \in \mathcal{V}$ and is associated with a C-dimensional feature vector $\mathbf{x}(v)$ extracted from the imaging data. The edges $\mathcal{E}$ of the graph represent the similarity between the subjects and incorporate the phenotypic information. The graph labelling is done in a semi-supervised fashion, through the use of a GCN trained on a subset of labelled graph vertices. Intuitively, label information will be propagated over the graph under the assumption that nodes connected with high edge weights are more comparable. An overview of the method is available in Fig. 1.

2.1   Databases and Preprocessing

We apply our model on two large and challenging databases for binary classification tasks. With the *ABIDE database*, we aim to separate healthy controls from ASD patients and exploit the acquisition information which can strongly affect the comparability of subjects. Our goal on the *ADNI database* is to predict whether an MCI patient will convert to AD. Our objective is to demonstrate the importance of exploiting longitudinal information, which can be easily integrated into our graph structure, to increase performance.

The **ABIDE database** [6] aggregates data from different acquisition sites and openly shares functional MRI and phenotypic data of 1112 subjects[1]. We select the same set of 871 subjects used in [1], comprising 403 individuals with ASD and 468 healthy controls acquired at 20 different sites. To ensure a fair comparison with the state of the art [1], we use the same preprocessing pipeline [3], which involves skull striping, slice timing correction, motion correction, global mean intensity normalisation, nuisance signal regression, band-pass filtering (0.01-0.1Hz) and registration of the functional MRI images to MNI152 standard anatomical space. The mean time series for a set of regions extracted from the Harvard Oxford (HO) atlas [5] were computed and normalised to zero mean and unit variance. The individual connectivity matrices $S_1, ..., S_N$ are estimated by computing the Fisher transformed Pearson's correlation coefficient between the representative rs-fMRI timeseries of each ROI in the HO atlas.

The **ADNI database** is the result of efforts from several academic and private co-investigators [2]. To date, ADNI in its three studies (ADNI-1, -GO and -2) has recruited over 1700 adults, aged between 55 and 90 years, from over 50 sites from the U.S. and Canada. In this work, a subset of 540 early/late MCI subjects that contained longitudinal T1 MR images and their respective anatomical segmentations was used. In total, 1675 samples were available, with 289 subjects (843 samples) diagnosed as AD at any time during follow-up and labelled as converters. Longitudinal information ranged from 6 to 96 months, depending on each subject. Acquisitions after conversion to AD were not included. As of 1st of July 2016 the ADNI repository contained 7128 longitudinal T1 MR images from 1723 subjects. ADNI-2 is an ongoing study and therefore data is still growing. Therefore, at the time of a large scale segmentation analysis (into 138 anatomical structures using MALP-EM [9]) only a subset of 1674 subjects (5074 images) was processed, from which the subset used here was selected.

2.2   Population graph construction

The proposed model requires two critical design choices: 1) the definition of the feature vector $\mathbf{x}(v)$ describing each sample, and 2) modelling the interactions between samples via the definition of the graph edges $\mathcal{E}$. We keep the feature vectors simple so as to focus on evaluating the impact of integrating contextual information in the classification performance. For the ABIDE data-set, we follow the method adopted by [1] and define a subject's feature vector as its vectorised functional connectivity matrix. Due to the high dimensionality of the connectivity matrix, a ridge classifier is employed to select the most discriminative features from the training set. For the ADNI dataset, we simply use the volumes of all 138 segmented brain structures.

The definition of the graph's edges is critical in order to capture the underlying structure of the data and explain the similarities between the feature vectors. We construct our sparse graph aiming to incorporate phenotypic information in

---

[1] http://preprocessed-connectomes-project.org/abide/
[2] http://adni.loni.usc.edu

our model, providing additional information on how similar two samples' feature vectors and labels are expected to be. Considering a set of H non-imaging measures $\mathbf{M} = \{M_h\}$ (e.g. subject's gender and age), the population graph's adjacency matrix W is defined as follows:

$$W(v, w) = Sim(S_v, S_w) \sum_{h=1}^{H} \rho(M_h(v), M_h(w)), \tag{1}$$

where, $Sim(S_v, S_w)$ is a measure of similarity between subjects, increasing the weights between the most similar graph nodes; ρ is a measure of distance between phenotypic measures. Considering categorical data such as gender or acquisition site, we define ρ as the Kronecker delta function δ. For quantitative measures such as the subject's age, we define ρ as a unit-step function with respect to a threshold θ: $\rho(M_h(v), M_h(w)) = \begin{cases} 1 & \text{if } |M_h(v) - M_h(w)| < \theta \\ 0 & \text{otherwise} \end{cases}$

The underlying idea behind this formulation is that non-imaging complementary data can provide key information explaining correlations between subjects' feature vectors. The objective is to leverage this information, so as to define an accurate neighbourhood system that optimises the performance of the subsequent graph convolutions. For the ABIDE population graph, we use $H = 2$ non-imaging measures, namely subject's *gender* and *acquisition site*. We define $Sim(S_v, S_w)$ as the correlation distance between the subjects' rs-fMRI connectivity networks after feature selection, as a separation between ASD and controls can be observed within certain sites. The main idea behind this graph structure is to leverage the site information, as we expect subjects to be more comparable within the same site due to the different acquisition protocols. The ADNI graph is built using the subject's *gender* and *age* information. These values are chosen because our feature vector comprises brain volumes, which can strongly be affected by age and gender. The most important aspect of this graph is the $Sim(S_v, S_w)$ function, designed to leverage the fact that longitudinal acquisitions from the same subject are present in the database. While linear classifiers treat each entry independently, here we define $Sim(S_v, S_w) = \lambda$ with $\lambda > 1$ if two samples correspond to the same subject, and $Sim(S_v, S_w) = 1$ otherwise, indicating the strong similarity between acquisitions of the same subject.

2.3   Graph Labelling using Graph Convolutional Neural Networks

Discretised convolutions, those commonly used in computer vision, cannot be easily generalised to the graph setting, since these operators are only defined for regular grids, e.g. 2D or 3D images. Therefore, the definition of localised graph filters is critical for the generalisation of CNNs to irregular graphs. This can be achieved by formulating CNNs in terms of spectral graph theory, building on tools provided by graph signal processing (GSP) [13].

The concept of spectral graph convolutions exploits the fact that convolutions are multiplications in the Fourier domain. The graph Fourier transform is defined by analogy to the Euclidean domain from the eigenfunctions of the Laplace

operator. The normalised graph Laplacian of a weighted graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, W\}$ is defined as $\mathcal{L} = I_N - D^{-1/2}WD^{-1/2}$ where I_N and D are respectively the identity and diagonal degree matrices. Its eigendecomposition, $\mathcal{L} = U\Lambda U^T$, gives a set of orthonormal eigenvectors $U \in \mathbb{R}^{N \times N}$ with associated real, non-negative eigenvalues $\Lambda \in \mathbb{R}^{N \times N}$. The eigenvectors associated with low frequencies/eigenvalues vary slowly across the graph, meaning that vertices connected by an edge of large weight have similar values in the corresponding locations of these eigenvectors.

The graph Fourier Transform (GFT) of a spatial signal $\mathbf{x}$ is defined on the graph $\mathcal{G}$ as $\hat{\mathbf{x}} \doteq U^T\mathbf{x} \in \mathbb{R}^N$, while the inverse transform is given by $\mathbf{x} \doteq U\hat{\mathbf{x}}$. Using the above formulations, spectral convolutions of the signal $\mathbf{x}$ with a filter $g_\theta = diag(\theta)$ are defined as $g_\theta * \mathbf{x} = g_\theta(\mathcal{L})\mathbf{x} = g_\theta(U\Lambda U^T)\mathbf{x} = Ug_\theta(\Lambda)U^T\mathbf{x}$, where $\theta \in \mathbb{R}^N$ is a vector of Fourier coefficients. Following the work of Defferrard et al. [4], we restrict the class of considered filters to polynomial filters $g_\theta(\Lambda) = \sum_{k=0}^{K} \theta_k\Lambda^k$. This approach has two main advantages: 1) it yields filters that are strictly localised in space (a K-order polynomial filter is strictly K-localised) and 2) it significantly reduces the computational complexity of the convolution operator. Indeed, such filters can be well approximated by a truncated expansion in terms of Chebyshev polynomials which can be computed recursively. Similarly to what is proposed in [8], we keep the structure of our GCN relatively simple. It consists of a series of convolutional layers, each followed by Rectified Linear Unit (ReLU) activation functions to increase non-linearity, and a convolutional output layer. The output layer is followed by a softmax activation function [8], while cross-entropy is used to calculate the training loss over all labelled examples. Unlabelled nodes are then assigned the labels maximising the softmax output.

3   Results

We evaluate our method on both the ADNI and ABIDE databases using a 10-fold stratified cross validation strategy. The use of 10-folds facilitates the comparison with the ABIDE state of the art [1] where a similar strategy is adopted. To provide a fair evaluation for ADNI, we ensure that the longitudinal acquisitions of the same subject are in the same fold (i.e. either the testing or training fold). We train a fully convolutional GCN with L hidden layers approximating the convolutions with $K = 3$ order Chebyshev polynomials. GCN parameters were optimised for each database with a grid search on the full database. For ABIDE, we use: $L = 1$, dropout rate: 0.3, l2 regularisation: 5.10^{-4}, learning rate: 0.005, number of features $C = 2000$. The parameters for ADNI are: $L = 5$, dropout rate: 0.02, l2 regularisation: 1.10^{-5}, learning rate: 0.01, graph construction variables $\lambda = 10$ and $\theta = 2$. The ABIDE network is trained for 150 epochs. Due to the larger network size, we train the ADNI network longer, for 200 epochs.

We compare our results to linear classification using a ridge classifier (using the scikit-learn library implementation [12]) which showed the best performance amongst linear classifiers. We investigate the importance of the population graph structure by using the same GCN framework with a random graph support of same density. Comparative boxplots across all folds between the three approaches

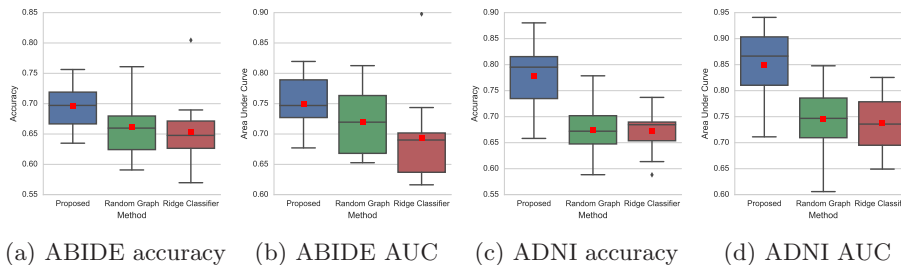(a) ABIDE accuracy   (b) ABIDE AUC   (c) ADNI accuracy   (d) ADNI AUC

Fig. 2: Comparative boxplots of the classification accuracy and area under curve (AUC) over all cross validation folds for the (a, b) ABIDE and (c, d) ADNI databases (MCI conversion task). The red dots correspond to the mean value.

are shown in Fig. 2 for both databases. GCN results (both with population and random graphs) are computed for ten different initialisation seeds and averaged. For both databases, we observe a significant ($p < 0.05$) increase both in terms of accuracy and area under curve using our proposed method, with respect to the competing methods. The random support yields equivalent or worse results to the linear classifier. For ABIDE, we obtain an average accuracy of 69.5%, outperforming the recent state of the art (66.8%) [1]. Results obtained for the ADNI database show a large increase in performance with respect to the competing methods, with an average accuracy of 77% on par with state of the art results [14], corresponding to a 10% increase over a standard linear classifier.

4   Discussion

In this paper, we introduced the novel concept of graph convolutions for population-based brain analysis. We proposed a strategy to construct a population graph combining image based patient-specific information with non-imaging based pairwise interactions, and use this structure to train a GCN for semi-supervised classification of populations. As a proof of concept, the method was tested on the challenging ABIDE and ADNI databases, respectively for ASD classification from a heterogeneous database and predicting MCI conversion from longitudinal information. Our experiments confirmed our initial hypothesis about the importance of contextual pairwise information for the classification process. In the proposed semi-supervised learning setting, conditioning the GCN on the adjacency matrix allows to learn representations even for the unlabelled nodes, thanks to the supervised loss gradient information that is distributed across the network. This has a clear impact on the quality of the predictions, leading to about 4.1% improvement for ABIDE and 10% for ADNI when comparing to a standard linear classifier (where only individual features are considered).

Several extensions could be considered for this work. Devising an effective strategy to construct the population graph is essential and far from obvious. Our

graph encompasses several types of information in the same edge. An interesting extension would be to use attributed graphs, where the edge between two nodes corresponds to a vector rather than a scalar. This would allow to exploit complementary information and weight the influence of some measures differently. Integrating time information with respect to the longitudinal data could also be considered. Our feature vectors are currently quite simple, as our main objective was to show the influence of the contextual information in the graph. We plan to evaluate our method using richer feature vectors, potentially via the use of autoencoders from MRI images and rs-fMRI connectivity networks.

References

1. Abraham, A., Milham, M., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G.: Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. NeuroImage (2016)
2. Brosch, T., Tam, R.: Manifold Learning of Brain MRIs by Deep Learning. In: MICCAI, pp. 633–640. Berlin, Heidelberg (2013)
3. Craddock, C., Sikka, S., Cheung, B., Khanuja, R., Ghosh, S., et al.: Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (C-PAC). Front Neuroinform 42 (2013)
4. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: NIPS. pp. 3837–3845 (2016)
5. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31(3), 968–980 (2006)
6. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., et al.: The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry 19(6), 659–667 (2014)
7. Havaei, M., Davy, A., et al: Brain tumor segmentation with Deep Neural Networks. Medical Image Analysis 35, 18–31 (Jan 2017)
8. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
9. Ledig, C., Heckemann, R.A., Hammers, A., et al.: Robust whole-brain segmentation: Application to traumatic brain injury. Med. Image Anal. 21(1), 40–58 (2015)
10. Niepert, M., Ahmed, M., Kutzkov, K.: Learning Convolutional Neural Networks for Graphs. arXiv preprint arXiv:1605.05273 (2016)
11. Parisot, S., Darlix, A., Baumann, C., Zouaoui, S., et al.: A Probabilistic Atlas of Diffuse WHO Grade II Glioma Locations in the Brain. PLOS ONE 11(1)
12. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., et al.: Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830 (2011)
13. Shuman, D.I., Narang, S.K., et al., P.: The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. 30(3), 83–98 (2013)
14. Tong, T., Gao, Q., Guerrero, R., Ledig, C., Chen, L., Rueckert, D.: A Novel Grading Biomarker for the Prediction of Conversion from Mild Cognitive Impairment to Alzheimer's Disease. IEEE Trans. Biomed. Eng. 64(1), 155–165 (2017)
15. Wolz, R., Aljabar, P., Hajnal, J.V., Lötjönen, J., Rueckert, D.: Nonlinear dimensionality reduction combining MR imaging with non-imaging information. Medical Image Analysis 16(4), 819–830 (May 2012)