Physica Medica 47 (2018) 103–111

Contents lists available at ScienceDirect

Physica Medica

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

Original paper

The effect of feature selection on multivariate pattern analysis of structural brain MR images

Ayşe Demirhan*

Gazi University, Ankara, Turkey

A R T I C L E I N F O

Keywords:
Brain MR images
Multivariate pattern analysis
Feature selection
Neurodegenerative disorders

A B S T R A C T

Clinical predictions performed using structural magnetic resonance (MR) images are crucial in neuroimaging studies and can be used as a successful complementary method for clinical decision making. Multivariate pattern analysis (MVPA) is a significant tool that helps correct predictions by exhibiting a compound relationship between disease-related features. In this study, the effectiveness of determining the most relevant features for MVPA of the brain MR images are examined using ReliefF and minimum Redundancy Maximum Relevance (mRMR) algorithms to predict the Alzheimer's disease (AD), schizophrenia, autism, and attention deficit and hyperactivity disorder (ADHD). Three state-of-the-art MVPA algorithms namely support vector machines (SVM), k-nearest neighbor (kNN) and backpropagation neural network (BP-NN) are employed to analyze the images from five different datasets that include 1390 subjects in total. Feature selection is performed on structural brain features such as volumes and thickness of anatomical structures and selected features are used to compare the effect of feature selection on different MVPA algorithms. Selecting the most relevant features for differentiating images of healthy controls from the diseased subjects using both ReliefF and mRMR methods significantly increased the performance. The most successful MVPA method was SVM for all classification tasks.

1. Introduction

Structural MR images provide a good quality view of the brain that can be used to describe the shape, size, and structures quantitatively. Improving the quality of images and developing new clinical diagnosis methods are active areas of brain MR imaging research [1–3]. Predicting neurodegenerative diseases using structural brain MR images is one of the fundamental purposes of neuroimaging studies where MVPA is used as a powerful tool. MVPA is beneficial where disease-related changes in the brain are subtle and spatially distributed that it is difficult to discriminate healthy and diseased images by using conventional mass-univariate methods like voxel-based morphometry. MVPA provides correction for multiple comparisons and statistical power for the prediction that improves its diagnostic value [4,5]. MVPA methods that use brain MR images are implemented successfully in previous clinical decision making studies as predictive tools to determine the clinical condition of the subjects [3–12].

Machine learning algorithms are employed frequently to evaluate multivariate patterns in the structural brain MR images for the purpose of classifying images as healthy or diseased for a number of neurodegenerative diseases [4,5,9,13]. Sabuncu and Konukoglu used SVM, the neighborhood approximation forest (NAF) and the relevance voxel machine (RVoxM) algorithms and common types of structural measurements from brain MR scans to predict an array of clinically relevant variables. Their results revealed that neurodegenerative diseases can be predicted from the brain MR images in a degree and MVPA produces better prediction accuracies than univariate models [4]. Ecker et al. investigated the predictive value of structural MR images in adults with autism using a whole-brain classification approach employing an SVM. They classified autism correctly at a specificity of 86.0% and a sensitivity of 88.0% [5]. Liu et al. utilized MVPA to classify major depressive disorder (MDD) patients with different therapeutic responses and healthy controls which combined searchlight algorithm and principal component analysis (PCA). According to the obtained results, they suggested that structural MR images with MVPA might be a useful and reliable method to study the neuroanatomical changes to differentiate patients with MDD from healthy controls [9]. Salvatore et al. analyzed T1-weighted MR images of 137 CE, 210 MCI and 162 healthy controls

* Address: Electrical and Electronics Engineering Department, Faculty of Technology, Gazi University, Beşevler, Teknikokullar, 06500 Ankara, Turkey.
E-mail address: ayseoguz@gazi.edu.tr.

https://doi.org/10.1016/j.ejmp.2018.03.002
Received 17 August 2017; Received in revised form 1 March 2018; Accepted 3 March 2018
1120-1797/

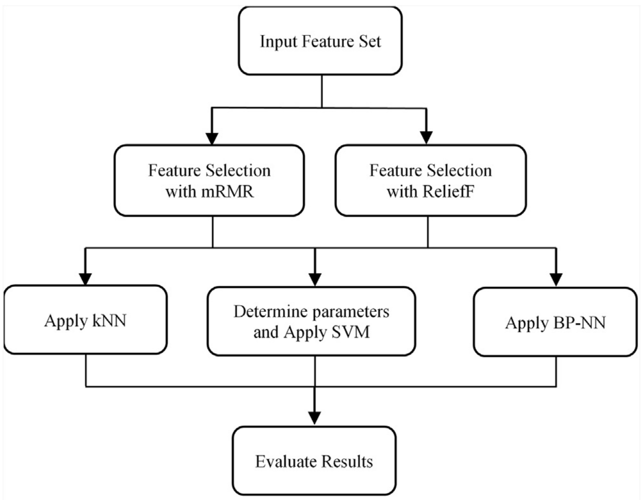

Fig. 1. Flow diagram of the system illustrating operating steps from top to bottom.

selected from the Alzheimer's disease neuroimaging initiative (ADNI) cohort to classify AD, MCI converters and MCI non-converters to AD. They selected the most discriminative features by PCA and used SVM for classification. Their classification accuracies were 76% for AD, 72% for MCI converters and 66% for MCI non-converters [13].

Feature selection is an essential operation to determine the effective subset of the input variables for a successful MVPA [7,14,15]. Input that is useful for classification has to be determined before MVPA analysis to ensure that it is meaningful in the condition of the disease and comparable across subjects. There are two types of feature selection methods namely feature ranking and feature subset selection. Feature ranking methods give a ranking score to each feature according to its degree of relevance that corresponds to the discriminative power of the feature for classification. Top-ranked features are then used for classification. Feature ranking methods are successful in high dimensional feature sets because of their good generalization ability. They have advantages like the independence of the classifier, lower computational cost and being fast. The disadvantage of these methods is that they do not have interaction with the classifier. Information gain, ReliefF, and mRMR can be listed as examples of feature ranking methods. Subset selection methods use a search strategy to determine a subset of features that jointly have discriminative power. Feature subset selection methods are not preferred for high dimensional problems since they are computationally expensive and have a risk of overfitting. Capturing feature dependencies is an advantage of these methods but selecting features depending on the classifier can be counted as a disadvantage. Correlation-based feature selection, consistency-based subset evaluation, and wrapper subset evaluation are some of the feature subset selection methods [14–17]. Two of the most frequently used feature ranking methods namely ReliefF and mRMR are used in this study for feature reduction because of their lower computational costs and independence of classifier since the same feature reduction method is applied to three different machine learning algorithms that are used to analyze images of 1390 people that belongs to four different neurodegenerative disease groups.

Previous neuroimaging studies that identify neurodegenerative diseases have proven that reducing the dimension of the input boost the classification accuracy and decrease the computation time by excluding the highly correlated features and features that are not valuable to discriminate between classes [18–22]. Demirhan et al. improved the accuracy of classifying AD and MCI using SVM up to 15% by selecting the most relevant features with ReliefF algorithm [18]. Cui et al. identified the conversion from MCI to AD by using mRMR method for feature selection to choose optimal subsets of features from each modality of data, then they employed the SVM by incrementally adding features based on their ranking till obtaining the highest area under the curve (AUC). They proved that the selected features are closely related to AD progression and verified the effectiveness of feature selection [19]. Wee et al. combined ranking and wrapper-based feature selection methods to identify the most relevant features for autism spectrum disorder classification. T-test and mRMR ranking based methods are used to reduce the number of features based on general characteristics of the data. Then SVM-based recursive feature elimination (SVM-RFE) is used to determine the subset of features. They obtained high classification accuracies up to 96% [20]. Castro et al. proposed a recursive feature elimination method that uses a machine learning algorithm based on composite kernels to the classification of healthy controls and patients with schizophrenia. They showed that feature selection improved the accuracy of classification and allowed a better identification of the brain regions that characterize schizophrenia [21]. Dai et al. integrated multimodal image features using multi-kernel learning and compared the effects of using different features for classification of ADHD patients. They selected optimal feature subset by combining feature ranking methods and feature subset selection methods. Their experiments showed that multi-kernel learning using selected multimodal features can yield better classification results for ADHD prediction [22].

In this study, MVPA analysis is performed to discriminate AD, schizophrenia, autism, and ADHD patients from the healthy controls using the morphometric features such as volumes and thickness of anatomical structures obtained from the T1-weighted structural brain MR images. Effect of using feature selection on the classification performance is investigated using ReliefF and mRMR feature ranking methods with an unbiased brain-wide approach. Three state-of-the-art machine learning algorithms, SVM, kNN, and BP-NN, are employed for the MVPA analysis. 5-fold cross validation (CV) is used for all feature selection and classification tasks to assess the generalization ability of

Table 1
Demographic features of the analyzed datasets that are constructed as age, sex, site-matched case and control groups.

Dataset	Variable	N per group	Age (Mean ± Std)		%Female	Number of sites
			Cases	Controls		
OASIS	AD	25	77.5 ± 6.8	77.5 ± 6.6	72	1
	AD mild	70	75.9 ± 7.3	76 ± 7.2	68.6	1
COBRE	Schizophrenia	50	34.3 ± 10.6	34.1 ± 10.7	18	1
MCIC	Schizophrenia	75	33.3 ± 11.6	33.4 ± 11.4	26.7	3
ABIDE	Autism	325	17.8 ± 7.4	17.9 ± 7.4	11.4	17
ADHD	ADHD	150	13.2 ± 2.4	13.2 ± 2.3	78.7	6