Neurocomputing 195 (2016) 112–116

Contents lists available at ScienceDirect

Neurocomputing

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

Multi-view ensemble learning for dementia diagnosis from neuroimaging: An artificial neural network approach

Jiajun Liu [a], Shuo Shang [b], Kai Zheng [c], Ji-Rong Wen [a,*]

[a] Renmin University of China, Beijing, China
[b] China University of Petroleum, China
[c] The University of Queensland, Australia

ARTICLE INFO

Article history:
Received 21 March 2015
Received in revised form
28 September 2015
Accepted 29 September 2015
Available online 4 February 2016

Keywords:
Artificial neural networks
Neuroimaging
Ensemble learning

ABSTRACT

Identifying abnormalities from neuroimaging of brain matters has been a crucial way of diagnosis of two closely associated diseases, namely Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI). Different types of neuroimaging have been developed to help such diagnosis, and significant research efforts are put into the automation and quantification of such diagnosis by computer algorithms over the past decades. In this paper we propose an ensemble learning framework to create effective models for AD/MCI related classification tasks from multiple modalities of neuroimaging and multiple baseline estimators. The framework is based on artificial neural networks and it resembles a composite model that solves the feature fusion learning problem as well as the prediction problem simultaneously, which targets at exploiting the prediction power of both fusing multiple data modalities and leveraging multiple mutually complementary classification models. We conduct extensive experiments on the well-known ADNI dataset and find that the proposed model works demonstrate advantages for both of the classification tasks studied.

1. Introduction

Dementia poses a serious challenge to the aging society. A study in 2005 [7] shows that 24.3 million people around the globe had been estimated to suffer from dementia of various types and degrees, and that an estimated 4.6 million people are newly diagnosed with dementia each year, at the time of the study. In particular, Alzheimer's Disease (AD), as the most common form of dementia among those over 65 years old, affects 26.6 million people, as reported by recent studies [24,6]. Brookmeyer et al. further predicted that by the year of 2050, 1.176% of the entire human population will be affected by AD [1], while developing countries such as China and India could even face a more dramatic growth (higher than 300%) of dementia occurrences by the time of 2040 [15]. It is also revealed in [2] that 6357 cases of AD out of all 9900 cases with dementia out of 254,367 subjects, marking AD a prevalent faction in all dementia occurrences. Earlier this decade, much research effort has been made to establish the links between Mild Cognitive Impairment (MCI) and AD [20,19], which subsequently suggests that an early diagnosis of MCI can help substantially for the identification of higher risks for AD [16,13].

Advances in computer vision/neuroimaging research and technologies have revealed that structural and functional features of the human brain may be significant changed by AD and MCI [10,11,26,23,12,14,33,32,30], providing neurohealth practitioners with new means for the timely and accurate diagnosis of AD and MCI. Prevalent medical imaging technologies include Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and CerebroSpinal Fluid (CSF). Each of those technologies represents a different perspective from which AD and MCI may be potentially diagnosed. MRI focuses on visualizing the structural brain atrophy, while PET can be used to identify the changes in the metabolism of the brain. CSF on the other hand captures the pathological amyloid depositions in the brain. Researchers have showed successful results for identifying AD and MCI from these imaging modalities [8,4,5,18]. Recently, progress has been achieved for using machine learning approaches to algorithmically estimate the risks of AD and MCI based on results from medical imaging. Many studies show that effective classification can be achieved for AD, MCI and Normal Control (NC), while some recent literature proposes to use regression models to predict "clinical scores" from MRI and PET data for dementia, such as Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) and Mini-Mental State Examination (MMSE) [3,9,21]. Efforts have also been made to create unified frameworks to combine the classification task and regression task [31].

* Corresponding author.
E-mail addresses: jiajunliu@ruc.edu.cn (J. Liu), jedi.shang@gmail.com (S. Shang), kevinz@itee.uq.edu.au (K. Zheng), jrwen@ruc.edu.cn (J.-R. Wen).

http://dx.doi.org/10.1016/j.neucom.2015.09.119
0925-2312/

Around many computer vision problems, researchers have also explored sparse feature selection/learning methods to exploit multiple feature modalities in order to improve the performance of machine learning models [29,28]. Meanwhile, though the idea of learning model ensembles for neuroimaging related tasks receives increasing attention [22,27], in general they are subject to particular models or kernel types, and are subject to a particular data modality. In this paper, we propose a generic, artificial neural network-based ensemble learning framework to harness the power of both multiple estimators as well as multiple views of data. The most appealing feature of this framework is its versatility and effectiveness – the framework can take any arbitrary set of data modalities as views, and take any arbitrary set of prediction models as estimators, and it can achieve very competitive classification results. Details of the proposed framework will be provided in Section 3.

The rest of the paper is organized as follows: first in Section 2, we describe the dataset we used to evaluate the proposed framework. Then in Section 3 we present the technical details of the proposed framework. Section 4 offers a comprehensive experimental evaluation on the proposed framework and its competitors. Finally we conclude the paper in Section 5.

2. Dataset

We use the publicly available ADNI (Alzheimer's Disease Neuroimaging Initiative) dataset for the evaluation of the proposed framework. Became operative in 2004, ADNI focuses on identifying AD in it earlier stages and tracking the progression of this disease by various biomarkers, which include samples of brain scans of various types. In this paper, we use the brain images from MRI and PET scans, in greyscale image format. For this dataset, ADNI operated a long-term study of AD progression on the involved subjects. Specifically, they followed approximately 200 individuals who were considered to have early AD for two years, 400 individuals with normal cognitive abilities for three years, and 200 individuals with MCI for three years. The evaluation of the proposed framework is conducted on a subset of 202 individuals as the baseline data, from which 51 are with AD, 99 are with MCI, and 52 are Normal Control (NC). The demographic information of the subjects is listed in Table 1. These scan images are preprocessed by ADNI for auto-correction for image quality and distortions.

3. Method

We mainly focus on the two binary classification tasks, i.e. AD vs. NC and MCI vs. NC. Unlike most of the existing methods which are often built on specific data modalities and prediction models, the proposed approach instead resembles a generic framework that can work with various types of modalities and classification models. The intuition behind it is that multiple modalities, or views, in the data, will introduce mutually-complementing distinguishing power for the application. On the other hand, multiple classification models enable us to harness the advantages of different forms of functions (linear, non-linear, kernel etc.). Combining the two layers of exploitation, we present an instantiation of the proposed framework in Fig. 1.

Table 1
Demographic information.

Group	AD	MCI	NC
No. subjects	51	99	52
Sex (F/M)	18/33	32/67	18/34
Avg. age	75.2	75.2	75.3
Avg. year of education	14.7	15.9	15.8

In this instance, the framework employs three views (e.g. MRI, PET and MRI+PET), and three estimators (e.g. linear regression, linear SVM, naive Bayes classifier). Thus the input layer consists of 9 nodes, with each node being the result of a specific combination of a view and an estimator. In Fig. 1, each rectangular area in the input layer represents an estimator, and each color represents a view. The views are first transformed into hidden features in the multi-view fusion layer, and then the results from all estimators are consolidated in the ensemble learning layer. The hidden units are then processed with a softmax layer, which then generates the prediction output. Next we provide the details of these layers. In the rest of the paper, the technical details of the neural network will be described mostly in vector forms, and we will use the assumptions and notations listed in Table 2.

3.1. Multi-view fusion

The multi-view fusion layer aims to harvest the results of an estimator in all views and sets to find feature representations from the individual results of the estimators in all views. In Fig. 1, this procedure is demonstrated as the sparse connections between the input layer and the fusion layer, which guarantees that the features learned from one estimator are using the results from that estimator only. We explain this with a case using one single sample. In mathematical form, the input for a sample (represents a training/test sample) is organized as a matrix $X_i \in \mathbb{R}^{n\_estimator \times n\_view}$

$$X_i = \begin{bmatrix} x_{e_1 v_1} & x_{e_1 v_2} & x_{e_1 v_3} & \cdots \\ x_{e_2 v_1} & x_{e_2 v_2} & x_{e_2 v_3} & \cdots \\ x_{e_3 v_1} & x_{e_3 v_2} & x_{e_3 v_3} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \quad (1)$$

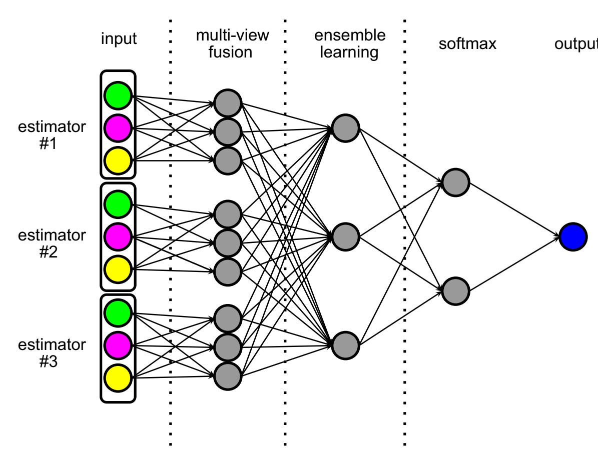

Fig. 1. An instance of the proposed framework with three estimators and three views. Nodes of the same color in the input layer indicate input from the same view, nodes in the same rectangular area indicate input from the same estimator. Note that bias nodes are not illustrated in this figure. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 2
Table of notations.

Notation	Description
$X \in \mathbb{R}^d$	The input of dimensionality d
$W^{(l)}$	The weights for the lth layer
$b^{(l)}$	The bias for the lth layer
$a^{(l)}$	The input of neurons in the lth layer
$z^{(l)}$	The intermediate values for the lth layer
$g^{(l)}(*)$	The activation function for the lth layer