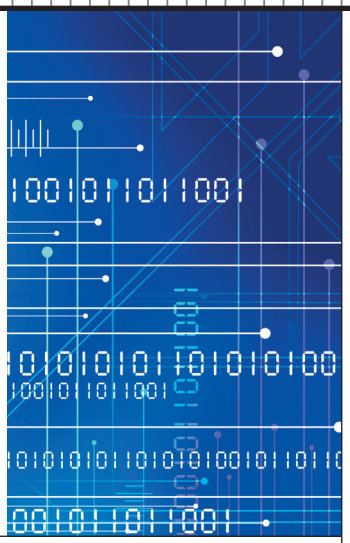
Machine Learning Approaches for the Neuroimaging Study of Alzheimer's Disease



Jieping Ye, Teresa Wu, and Jing Li, *Arizona State University*
Kewei Chen, *Banner Alzheimer's Institute and Shanghai Jiao Tong University*

Machine learning tools aid many Alzheimer's disease-related investigations by enabling multisource data fusion and biomarker identification as well as analysis of functional brain connectivity.

Alzheimer's disease (AD) is the most common type of dementia, accounting for 60-80 percent of age-related dementia cases. AD progressively destroys neurons and their connections in the brain, leading to loss of cognitive function and, ultimately, death.

The disease currently affects about 5.3 million people in the US, and the number of victims will significantly increase in the near future without the development of therapeutics. AD was the seventh-leading cause of death across all ages in the US in 2006; it was the fifth-leading cause of death for those 65 and older (www.alz.org). The direct cost to care for AD patients by family members or healthcare professionals is more than $100 billion per year; this figure is expected to rise dramatically as the population ages during the next several decades.

To avert a healthcare crisis, AD researchers have recently intensified their efforts to delay, cure, or prevent the onset and progression of the disease. These efforts have generated a large amount of data, including brain neuroimages, that provides unprecedented opportunities to investigate AD-related questions with higher confidence and precision. Especially promising is the use of machine learning approaches to analyze neuroimages to improve AD detection and diagnosis. Emerging techniques include the fusion of AD data from multiple sources, AD biomarker identification from multiple sources, and the analysis of functional brain connectivity.

NEUROIMAGING

Recent studies have demonstrated that imaging parameters based on brain scans are more sensitive and consistent measures of AD disease diagnosis and progression than cognitive assessment. Thus, neuroimaging techniques offer great potential to identify the specific biomarkers that can identify individuals early in the course of a dementing illness even before onset of the disease.

One common neuroimaging technique is structural magnetic resonance imaging (MRI), which visualizes brain anatomy with a high degree of contrast between brain tissue types. Researchers can use structural MRI to measure specific structures such as the hippocampus, entorhinal cortex, and amygdyla to detect abnormal volumetric changes related to AD. Another popular neuroimaging technique is positron emission tomography (PET). Using different radioactive tracers, PET provides information on various physiological, biochemical, and metabolic processes.

Recognizing these technologies' importance, the National Institutes of Health in 2004 funded the Alzheimer's Disease Neuroimaging Initiative (ADNI; www.adni-info.org), which has become a landmark study in the development of neuroimaging and other biosignatures for the disease. All ADNI subjects undergo 1.5T or 3T structural MRI scans (T, for Tesla, is a unit of magnetic flux density); half undergo fluorodeoxyglucose (FDG) PET scans.

FDG-PET scans have been shown to be highly sensitive at detecting AD-related glucose uptake abnormalities even before onset of the disease. For example, Eric Reiman and his colleagues examined the cerebral metabolic rate for glucose (CMRgl)

0018-9162/11/$26.00 © 2011 IEEE Published by the IEEE Computer Society

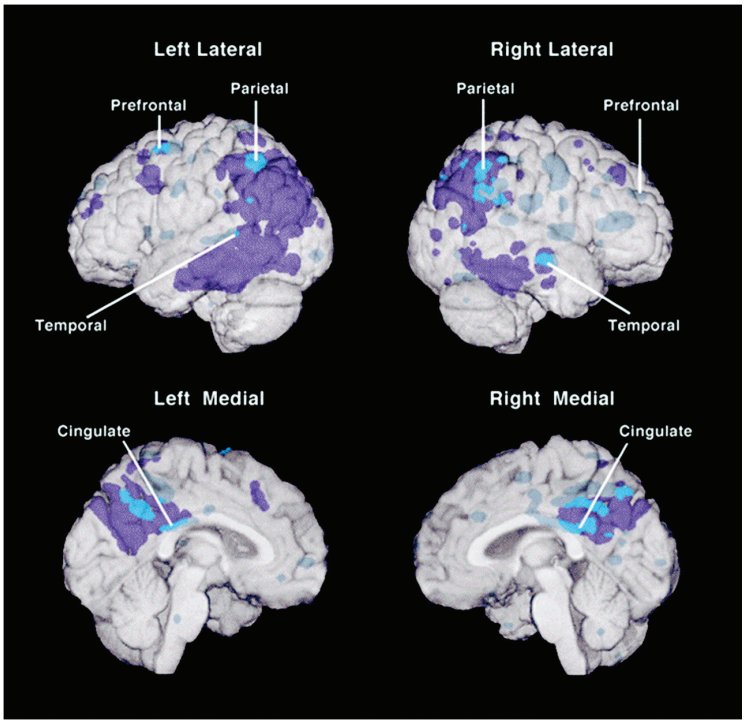


Figure 1. Brain regions with abnormally low CMRgl in young adult carriers of the ApoE4 allele and their relation to brain regions with abnormally low CMRgl in patients with probable AD. Source: E. Reiman et al., "Functional Brain Abnormalities in Young Adults at Genetic Risk for Late-Onset Alzheimer's Dementia," *Proc. National Academy of Sciences*, 6 Jan. 2004, pp. 284-289).

among normally healthy young subjects who had 0, 1, or 2 copies of the apolipoprotein E4 (ApoE4) allele, a known genetic risk factor associated with AD (E. Reiman et al., "Functional Brain Abnormalities in Young Adults at Genetic Risk for Late-Onset Alzheimer's Dementia," *Proc. National Academy of Sciences*, 6 Jan. 2004, pp. 284-289). As Figure 1 shows, the healthy ApoE4 carriers had lower CMRgl (bright blue) than the noncarriers in brain regions whereas CMRgl was abnormally low in AD patients (purple areas).

In addition to neuroimaging data, ADNI compiles demographic information such as age, gender, and years of education; genetic markers (such as possession of ApoE4 allele); clinical ratings; various cognitive tests; and protein abnormalities in cerebrospinal fluid (CSF), also associated with AD.

MULTISOURCE DATA FUSION

Current research has focused on using either regions of interest (ROIs) or a voxel-based approach to extract features from one neuroimaging modality—for example, structural MRI or PET alone. Integrating complementary ROI and voxel-based information from different neuroimaging sources and incorporating additional information such as demographic and genetic data will likely improve the sensitivity and specificity of AD detection.

One way to combine numerous AD data sources is to treat the variables in all datasets indiscriminately, without considering their different levels of relevance to AD. Toward this end, AD researchers have explored *multiple kernel learning* (MKL), a technique that synthesizes information from multiple heterogeneous data sources into a rigorous optimization problem (J. Ye et al., "Heterogeneous Data Fusion for Alzheimer's Disease Study," *Proc. 14th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, ACM Press, 2008, pp. 1025-1033). MKL works by first constructing a kernel from each of the data sources and then combining these kernels into a single one based on a certain criterion for improved classification performance.

To illustrate, assume we have a collection of p data sources with n samples. We first construct p kernel Gram matrices $\{G_i\}_{i=1,\ldots,p}$ of size n by n, one for each of the p data sources. A kernel Gram matrix computes the dot product of the samples in some feature space, thus capturing the similarities between samples. For example, the (j,k)-th entry of the matrix G_i captures the similarity between the j-th sample and the k-th sample based on the i-th data source alone. MKL integrates these p data sources by computing a composite kernel Gram matrix $G = \Sigma_i \theta_i G_i$, where the optimal combination coefficients are obtained by optimizing a certain criterion. Figure 2 illustrates the use of MKL to fuse data from five AD-related sources: structural MRI, PET, genetic, CSF, and demographic.

To separate AD patients from normal control (NC) subjects, researchers in the MKL study cited above applied the technique to fuse structural MRI data based on tensor factorization, structural MRI data based on anatomical automatic labeling, genetic information based on ApoE4, gender, and age. They measured performance in terms of sensitivity (accuracy in correctly identifying AD patients) and specificity (accuracy in correctly identifying NC subjects). MKL achieved 95.0 percent sensitivity and 89.5 percent specificity, significantly outperforming even the best prediction (80.0 percent sensitivity and 79.5 percent specificity) using each of the five data sources individually.

MULTISOURCE BIOMARKER SELECTION

Another urgent task in current AD research is biomarker identification, which can be considered a general feature selection problem. Feature selection algorithms attempt to remove as many irrelevant and redundant features as possible and to find a feature subset such that, with dimensionally reduced data, a learning algorithm can achieve better performance.

Feature selection has a wide variety of applications including text mining, image processing, and biomedical informatics. Traditional feature selection algorithms work on a single data source only. The challenge for AD researchers is developing effective algorithms for data from multiple sources to enable the identification of composite biomarkers.

MKL fuses the contributions of biomarkers from multiple data sources. Specifically, the combined kernel matrix extracts data patterns in the form of pairwise similarities, which can then serve as the input for a generic feature selection algorithm. Preliminary studies indicate that MKL not only adequately distinguishes AD and NC subjects but also identifies brain regions that play more significant roles than others in AD.

FUNCTIONAL BRAIN CONNECTIVITY

Recent studies have shown that higher cognitive processing results from different brain regions interacting with one another rather than working independently. Dramatic global cognitive decline is a major symptom of AD, and patients' brains thus exhibit abnormal patterns of functional connectivity—for example, reduced hippocampal network activity.

Some studies of early AD and mild cognitive impairment have found increased connectivity between the frontal lobe and other brain regions. Researchers have interpreted this as a compensatory reallocation or recruitment of cognitive resources. Because regions in the frontal lobe are typically affected later in the course of the disease, an increase in frontal connectivity arguably helps early AD patients maintain some memory and attention abilities.

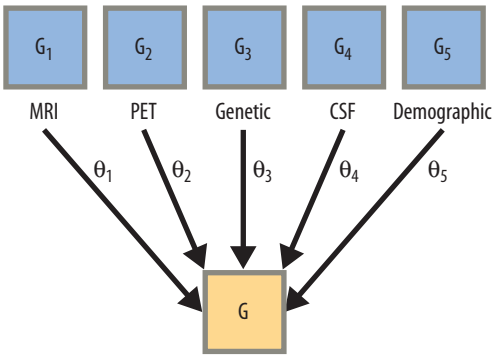


Figure 2. Illustration of using multiple kernel learning to fuse data from five sources: structural magnetic resonance imaging (MRI), positron emission tomography (PET), genetic, cerebrospinal fluid (CSF), and demographic. The composite kernel matrix G is a linear combination of the five kernel matrices constructed from these five data sources.

A method based on *sparse inverse covariance estimation* (SICE) identifies functional brain connectivity networks from FDG-PET data (S. Huang et al., "Learning Brain Connectivity of Alzheimer's Disease by Sparse Inverse Covariance Estimation," *NeuroImage*, 15 Apr. 2010, pp. 935-949). SICE makes it possible to identify both the connectivity network structure and connectivity strength for a large number of brain regions with small sample sizes.

Researchers using SICE have observed distinct connectivity patterns for AD patients and NC subjects in terms of the number of connections within lobes, between lobes, and between hemispheres, and in terms of the strength of such connections—findings consistent with the AD literature. SICE techniques can also identify connectivity-based FDG-PET biomarkers.

Machine learning tools aid many AD-related investigations by enabling multisource data fusion and biomarker identification as well as analysis of functional brain connectivity. Despite these advances, many challenges remain, including more effectively predicting disease progression and using multisource data for efficient clinical treatment evaluation.

Jieping Ye *is an associate professor of computer science, as well as a core faculty member of the Center for Evolutionary Medicine and Informatics, at Arizona State University (ASU). Contact him at jieping.ye@asu.edu.*

Teresa Wu *is an associate professor of industrial engineering at ASU. Contact her at teresa.wu@asu.edu.*

Jing Li *is an assistant professor of industrial engineering at ASU. Contact her at jing.li.8@asu.edu.*

Kewei Chen *is a senior scientist at Banner Alzheimer's Institute as well as director of the Computational Image Analysis lab at Banner Good Samaritan Medical Center's PET Center; he is also an adjunct professor at Med-X Research Institute, Shanghai Jiao Tong University, China. Contact him at kewei.chen@bannerhealth.com.*

Editor: Naren Ramakrishnan, Dept. of Computer Science, Virginia Tech, Blacksburg, VA; naren@cs.vt.edu