CDT Journal Club - First Written Review

Automatic classification of MR scans in Alzheimer's disease

Stefan Klppel , Cynthia M. Stonnington , Carlton Chu , Bogdan Draganski , Rachael I. Scahill , Jonathan D. Rohrer , Nick C. Fox , Clifford R. Jack , John Ashburner , Richard S. J. Frackowiak

Review Author:
Răzvan Valentin Marinescu
razvan.marinescu.14@ucl.ac.uk

Paper chosen by:
Prof. Sebastien Ourselin

EPSRC Centre for Doctoral Training in Medical Imaging
University College London

February 19, 2015

Paper background

Dementia and associated diseases such as Alzheimer's disease (AD) affect 36 million people globally [1] and account for 486,000 deaths [2]. Definitive diagnosis of AD can only be made with histopathological confirmation of amyloid plaques and neurofibrillary tangles, normally at autopsy. Early accurate detection of AD is important because the treatment is most effective if undertaken as early as possible. Medical imaging modalities such as MRI have been used for automatic AD diagnosis by measuring the atrophy rates of cortical volumes [3, 4, 5], but these methods have not yet been introduced in clinical practice.

Methods

This paper uses a machine-learning classifier called *Support Vector Machine* (SVM) for the purpose of AD diagnosis. It is the first study to use SVMs for diagnosing AD using pathologically confirmed cases. The authors futher use SVMs to differentially diagnose AD and frontotemporal lobal degeneration (FTLD). The study uses 4 different cohorts of patients: (I and II) the first two sets consisted of neuropathologically-confirmed AD subjects, (III) the third group consisted of probable mild-AD patients limited to 80 years of age or younger and (IV) the fourth set was made of subjects with neuropathologically proven FTLD having comparable Mini Mental State Examination (MMSE) scores with the first two groups. The scans were collected over a period up to 10 years using between 3-13 different scanners depending on the cohort.

Several image processing steps were taken in order to remove artefacts and ultimatelly register the scans. After visual inspection for structural abnormalities, images were segmented into grey matter (GM), white matter (WM) and cerebro-spinal fluid (CSF) using SPM5 [6]. The GM segments were further normalised to the population templates using a *diffeomorphic registration algorithm* [7], which was a sophisticated technique at the time the paper was published. In order to ensure the overall ammount of each tissue class remained constant, a separate *'modulation'* step was undertaken as described by Ashburner and Friston. [8]

Classification of patients was performed using Support-Vector Machine (SVM) which is a method of supervised, binary classification where voxels from MR images are treated as points in a high-dimensional space. [9, 10] Once projected in this higher-dimensional space, the SVM tries to find the *optimal separating hyperplane* (OSH) that separates the two classes as efficient as possible. The OSH is defined by the voxels that are the closest to the separating boundary, called *support vectors*. The OSH, wich is produced during training, can then be used to classify a new patient by finding on which side of the OSH it's corresponding data point lies.

Key results

Klöppel et al. [11] show that the linear SVM is able to diagnose AD vs controls and AD vs FTLD with a high accuracy (95%, 92.9% and 81.1% for cohorts I, II and III). This is in line with the diagnosis rates of the best clinical centers around the world. [12] If using the antero-medial lobe volume ROI, accuracy improves only for mild-AD (88.9%). Authors have also tried to use different datasets for training and testing, and again the observed accuracy is very high (87.5% and 96.4%), suggesting that the method is robust and can be generalised across multiple scanners.

A nice feature of the SVM is that it allows the localisation of the voxels that were relevant for classification (i.e. the support vectors). Authors show that for AD vs controls, most of these vectors were clustered around the parahyppocampal gyrus and parietal cortex. They also tried to use non-linear kernel functions but this failed to improve performance. Overall, the results they obtained were equally good or better than other methods used in the field based on MR images. [13, 14, 4, 15, 5] The work of Klöppel et al. [11] shows that automated, data-driven

techniques are equally good or even better than clinical experts at diagnosing AD patients, because NINCDS-ADRDA or DSM-IIIR criteria have an average sensitivity and specificity of 81% and 70% respectively [12], which are lower than the values reported in the paper.

Limitations

One methodological limitation is that the SVM is a binary classifier, where only two classes can be distinguished at the same time. This can be overcome using a multiclass-SVM, which could reduce the original multi-class problem to several binary problems [16] or use directed acyclic graphs [17]. However, other multiclass classifiers might provide a more natural solution such as logistic regression [18] or a parzen-window classifier with a linear kernel. [19] Moreover, the method of Klöppel [11] assumes there is no head motion during the scan and does not correct it, so future studies can incorporate a motion-correction algorithm. Another limitation is that the diffeomorphic registration technique by Ashburner, 2007 [7] which is based on sums of squares, is very sensitive to affine initialization. [20] One factor that limits the applicability of the method to clinical practice is that a user is required to check for registration errors. This can be overcome by alerting the user only when a certain threshold is reached and by improving the segmentation method.

Impact and conclusion

The work of Klöppel et al. [11] is the first to perform automatic classification of MRI scans in pathologically-confirmed AD and to differentiate different forms of dementia. It clearly made a huge impact when it was published in 2008 and so far got 569 citations. Although it is a methodological paper, it managed to get published in a clinical journal like Brain probably due to the unique datasets it analysed that contained pathologically-confirmed AD patients. Since it was published, the field has evolved and now people look at diagnosing various subtypes of AD or MCI [21] or other types of dementias such as schizophrenia [22], Parkinson syndrome [23] or multiple systems atrophy. [23] Others like Zhang et al. [24] have applied SVMs to multi-modal data combining structural MR imaging, functional imaging (FDG-PET) and CSF biomarkers to classify AD patients. Later in 2008, Klöppel et al. [25] did a more detailed study comparing the diagnostic accuracy between the SVM method and radiologists and found SVMs to have similar results to expert radiologists, the advantage of SVMs being that they don't require expert-knowledge and can easily be exchanged between centers.

In conclusion, this study was fundamental in the field of computer-aided diagnosis of neurological disorders showing that SVMs can successfully diagnose AD from healthy aging subjects. Several modifications are needed in order to make it applicable in a clinical environment such as shorter computation time or making the segmentation fully automated without having a user check the intermediate results. Nevertheless, it highlights the increasingly important role that computer-aided diagnosis of dementia will have in clinical practice in the near future.

Bibliography

[1] World Health Organization et al. Dementia: Fact sheet no. 362. *Retrieved June*, 30:2012, 2012.

[2] Rafael Lozano, Mohsen Naghavi, Kyle Foreman, Stephen Lim, Kenji Shibuya, Victor Aboyans, Jerry Abraham, Timothy Adair, Rakesh Aggarwal, Stephanie Y Ahn, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. *The Lancet*, 380(9859):2095–2128, 2013.

[3] Nick C Fox and Jonathan M Schott. Imaging cerebral atrophy: normal ageing to alzheimer's disease. *The Lancet*, 363(9406):392–394, 2004.

[4] Josephine Barnes, Rachael I Scahill, Richard G Boyes, Chris Frost, Emma B Lewis, Charlotte L Rossor, Martin N Rossor, and Nick C Fox. Differentiating ad from aging using semiautomated measurement of hippocampal atrophy rates. *Neuroimage*, 23(2):574–581, 2004.

[5] Lars-Olof Wahlund, Ove Almkvist, Kaj Blennow, Knut Engedahl, Aki Johansson, Gunhild Waldemar, and Henrike Wolf. Evidence-based evaluation of magnetic resonance imaging as a diagnostic tool in dementia workup. *Topics in Magnetic Resonance Imaging*, 16(6):427–437, 2005.

[6] SPM5; Wellcome Trust Centre for Neuroimaging, London, UK. `http://www.fil.ion.ucl.ac.uk/spm`.

[7] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.

[8] John Ashburner and Karl J Friston. Voxel-based morphometrythe methods. *Neuroimage*, 11(6):805–821, 2000.

[9] Vladimir Naumovich Vapnik and Vlamimir Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998.

[10] Christopher M Bishop et al. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[11] Stefan Klöppel, Cynthia M Stonnington, Carlton Chu, Bogdan Draganski, Rachael I Scahill, Jonathan D Rohrer, Nick C Fox, Clifford R Jack, John Ashburner, and Richard SJ Frackowiak. Automatic classification of mr scans in alzheimer's disease. *Brain*, 131(3):681–689, 2008.

[12] DS Knopman, St T DeKosky, JL Cummings, H Chui, J Corey-Bloom, N Relkin, GW Small, B Miller, and JC Stevens. Practice parameter: Diagnosis of dementia (an evidence-based review) report of the quality standards subcommittee of the american academy of neurology. *Neurology*, 56(9):1143–1153, 2001.

[13] KM Gosche, JA Mortimer, CD Smith, WR Markesbery, and DA Snowdon. Hippocampal volume as an index of alzheimer neuropathology findings from the nun study. *Neurology*, 58(10):1476–1482, 2002.

[14] CR Jack, DW Dickson, JE Parisi, YC Xu, RH Cha, PC Obrien, SD Edland, GE Smith, BF Boeve, EG Tangalos, et al. Antemortem mri findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology*, 58(5):750–757, 2002.

[15] John G Csernansky, Julia Hamstra, Lei Wang, Daniel McKeel, Joseph L Price, Mokhtar Gado, and John C Morris. Correlations between antemortem hippocampal volume and postmortem neuropathology in ad subjects. *Alzheimer Disease & Associated Disorders*, 18(4):190–195, 2004.

[16] Kai-Bo Duan and S Sathiya Keerthi. Which is the best multiclass svm method? an empirical study. In *Multiple Classifier Systems*, pages 278–285. Springer, 2005.

[17] John C Platt, Nello Cristianini, and John Shawe-Taylor. Large margin dags for multiclass classification. In *nips*, volume 12, pages 547–553, 1999.

[18] Rahul S Desikan, Howard J Cabral, Christopher P Hess, William P Dillon, Christine M Glastonbury, Michael W Weiner, Nicholas J Schmansky, Douglas N Greve, David H Salat, Randy L Buckner, et al. Automated mri measures identify individuals with mild cognitive impairment and alzheimer's disease. *Brain*, 132(8):2048–2057, 2009.

[19] John Shawe-Taylor and Nello Cristianini. *Kernel methods for pattern analysis*. Cambridge university press, 2004.

[20] Brian B Avants, Nicholas J Tustison, Gang Song, Philip A Cook, Arno Klein, and James C Gee. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044, 2011.

[21] Sven Haller, Pascal Missonnier, FR Herrmann, Cristelle Rodriguez, M-P Deiber, Duy Nguyen, Gabriel Gold, K-O Lovblad, and Panteleimon Giannakopoulos. Individual classification of mild cognitive impairment subtypes by support vector machine analysis of white matter dti. *American Journal of Neuroradiology*, 34(2):283–291, 2013.

[22] Babak A Ardekani, Ali Tabesh, Serge Sevy, Delbert G Robinson, Robert M Bilder, and Philip R Szeszko. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Human brain mapping*, 32(1):1–9, 2011.

[23] Niels K Focke, Gunther Helms, Sebstian Scheewe, Pia M Pantel, Cornelius G Bachmann, Peter Dechent, Jens Ebentheuer, Alexander Mohr, Walter Paulus, and Claudia Trenkwalder. Individual voxel-based subtype prediction can differentiate progressive supranuclear palsy from idiopathic parkinson syndrome and healthy controls. *Human brain mapping*, 32(11):1905–1915, 2011.

[24] Daoqiang Zhang, Yaping Wang, Luping Zhou, Hong Yuan, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative, et al. Multimodal classification of alzheimer's disease and mild cognitive impairment. *Neuroimage*, 55(3):856–867, 2011.

[25] Stefan Klöppel, Cynthia M Stonnington, Josephine Barnes, Frederick Chen, Carlton Chu, Catriona D Good, Irina Mader, L Anne Mitchell, Ameet C Patel, Catherine C Roberts, et al. Accuracy of dementia diagnosisa direct comparison between radiologists and a computerized method. *Brain*, 131(11):2969–2974, 2008.