

A novel dementia diagnosis strategy on arterial spin labeling magnetic resonance images via pixel-wise partial volume correction and ranking

Wei Huang · Peng Zhang · Minmin Shen

Received: 11 July 2014 / Revised: 9 October 2014 / Accepted: 24 November 2014 /
Published online: 6 December 2014
© Springer Science+Business Media New York 2014

Abstract Arterial Spin Labeling (ASL) is an emerging magnetic resonance imaging technique attracting increasing attention in dementia diagnosis only beginning from recent years. ASL is capable to provide direct and quantitative measurement of cerebral blood flow (CBF) of scanned patients, so that brain atrophy of demented patients could be revealed by measured low CBF within certain brain regions through ASL. However, partial volume effects (PVE) mainly caused by signal cross-contamination due to pixel heterogeneity and limited spatial resolution of ASL, often prevents CBF from being precisely measured. Inaccurate CBF is prone to mislead and even deteriorate dementia disease diagnosis results, thereafter. In this paper, a novel dementia disease diagnosis strategy based on ASL is proposed for the first time. The diagnosis strategy is composed of two steps: 1) to conduct pixel-wise PVE correction on original ASL images and 2) to predict dementia disease severities based on corrected ASL images via ranking. Extensive experiments and comprehensive statistical analysis are carried out to demonstrate the superiority of the new strategy with comparison to several existing ones. Promising results are reported from the statistical point of view.

Keywords Magnetic resonance image · Alzheimer's disease · Ranking

W. Huang (✉)

School of Information Engineering, Nanchang University, Nanchang, China
e-mail: huangwei@ncu.edu.cn

P. Zhang (✉)

School of Computer Science, Northwestern Polytechnical University, Xi'an, China
e-mail: zh0036ng@nwpu.edu.cn

M. Shen

INCIDE Center, University of Konstanz, Konstanz, Germany
e-mail: minmin.shen@uni-konstanz.de

M. Shen

School of Software Engineering, South China University of Technology, Guangzhou, China

1 Introduction

Alzheimer's Disease (AD), the most common form of dementia, is often diagnosed in patients over 60 years old and generally regarded as one of the five most severe non-communicable diseases worldwide reported by the World Health Organization (i.e. others include cardiovascular disease, cancer, diabetes and chronic lung disease) [33]. According to a recent population study conducted by the United Nations, there are already over 26.6 million AD patients diagnosed globally [30], and 1 in 85 people of the whole world is predicted to be suffering from AD by the year 2050 [3]. It is also widely acknowledged that, accurate diagnosis and timely treatment is essential to delay the onset and progression of AD [3].

In order to accurately diagnose the progression of dementia diseases, a variety of methods have been proposed and utilized to date. Popular diagnosis methods include pathography analysis, cognitive examination, and brain scanning. Pathography [21] is helpful to predict curable symptoms of demented patients, who may usually suffer from other forms of diseases (e.g., stroke, heart disease, renal failure, etc) simultaneously. Cognitive examination evaluates the progression of demented patients through a series of cognition tests based on diverse cognitive capabilities of patients, including short-memory, long-memory, logic analysis, etc. Popular cognitive examinations include Mini-Mental State Examination (MMSE) [8] and Addenbrooke's Cognitive Examination (ACE) [23]. Although these cognitive exams are easy to be carried out by clinicians, their outcomes could be highly biased by patients specialities. For instance, highly educated patients with dementia disease are likely to outperform non-educated ordinary patients without dementia in those cognitive exams. For brain scanning, it is widely accepted as an effective and affordable way in dementia diagnosis nowadays, and Magnetic Resonance Imaging (MRI) is generally regarded as a powerful scanning tool and receives vast popularity because it is free of ionizing radiation exposure, compared with other scanning tools such as Computed Tomography (CT) and Positron Emission Tomography (PET), for patients safety consideration. Different MRI scanning techniques, including both structural MRI (sMRI) and functional MRI (fMRI), have already been widely incorporated in clinical dementia diagnosis at present [1, 4–6, 20, 22, 24, 25, 27, 31, 32, 35].

Arterial Spin Labeling (ASL) is an emerging perfusion fMRI technique attracting increasing attention in dementia studies only beginning from recent years [22, 25]. Compared with other popular perfusion MRI techniques such as Dynamic Contrast Enhanced MRI (DCE-MRI) or Dynamic Susceptibility Contrast MRI (DSC-MRI), ASL requires no injection of external contrast enhancement agent (e.g., gadolinium) into patients while being scanned. Thus, anaphylactoid reactions [24] unexpectedly caused by those agents on certain patients could be totally avoided for ASL. Technically, an ASL-MRI image is produced by two types of images: a label image and a control image. Their acquisition steps are illustrated in Fig. 1. The yellow region 2 in Fig. 1a and the green region 4 in Fig. 1b describe the same Region-of-Interest (ROI), in which ASL-MRI images are acquired. The purple region 1 in Fig. 1a represents an area where arterial blood water is magnetically labeled via a 180 degree Radio-Frequency (RF) inversion pulse. In this way, water molecules within the arterial blood are magnetically labeled and utilized as the "tracer", instead of the conventional injected contrast enhancement agent. Label images are taken when labeled blood water flows into the ROI, and example label images from the transverse view acquired from one patient in this study are displayed in Fig. 1a. For control images, the blood water is not magnetically labeled, and control images are taken at the same ROI directly. Example control images of the same patient are displayed in Fig. 1b. Although label and control images

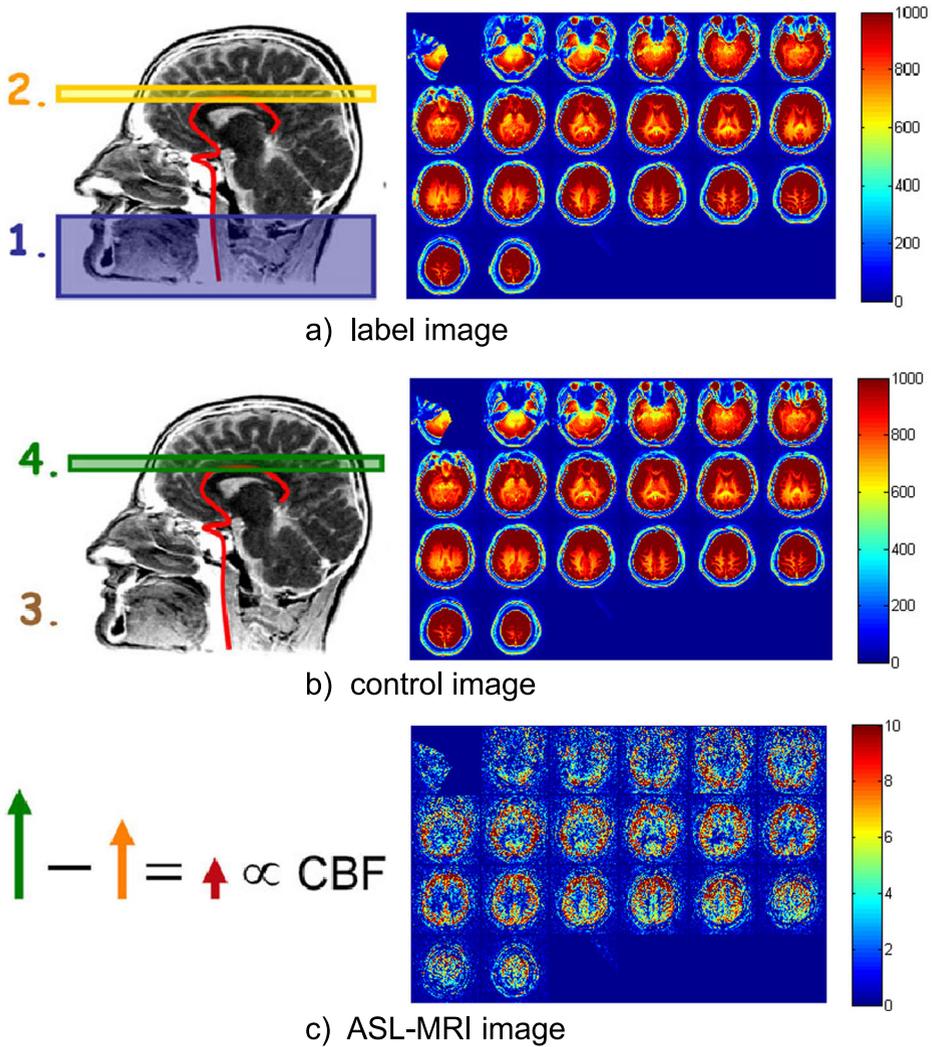


Fig. 1 An illustration of ASL-MRI acquisition with example images displayed from the transverse view (the plotting scale unit is [mL/(100g*min)])

look similar towards each other, difference exists between them and ASL-MRI are produced as the difference (i.e., using a control image minus a corresponding label image), and example ASL-MRI of the same patient is illustrated in Fig. 1c. Generally speaking, the *Cerebral Blood Flow* (CBF) on each pixel of ASL-MRI is proportional towards its ASL signal, and brain atrophy within certain brain regions of demented patients can be revealed by low CBF within those regions, compared with ones of ordinary people, reflected in ASL-MRI.

Although ASL is a promising bio-marker for disease diagnosis and progression analysis in dementia, the problem of *Partial Volume Effect* (PVE) should be carefully tackled. PVE is generally defined as the loss of apparent activity in small objects because of the limited resolution of an imaging system [27]. In ASL, since its spatial resolution is not high

(i.e., it can be perceived by example images in Fig. 1), pixels in ASL-MRI images containing various tissues of *Gray Matter* (GM), *White Matter* (WM) and *Cerebro-Spinal Fluid* (CSF) are likely to be assigned with under-estimated ASL signal and low CBF quantities, which reflects the loss of apparent activity in ASL-MRI because of PVE. In order to correct PVE, there are already several studies conducted in recent years [1, 5, 6], and the regression-based method receives much popularity among them [1]. However, its shortcoming is also obvious. Neighboring pixels are usually indispensable for PVE correction on each single pixel of ASL-MRI, making blurring and loss of brain details inevitable in correction results of this method [1]. A case in point is illustrated in the 1st row of Fig. 2. Therefore, in order to enable ASL a reliable indicator for the following dementia diagnosis, the problem of PVE needs to be properly handled first.

After PVE correction on ASL-MRI is conducted, the next critical step in dementia diagnosis is to predict the dementia disease severity based on corrected ASL-MRI of each patient. Dementia studies incorporating ASL-MRI only begin to emerge in recent years [4, 22, 25, 31], and most of them concentrate on verifying ASL-MRI as a new indicator in identifying dementia disease, with comparison to other previously well-established imaging modalities (e.g. PET [25] and FDG-PET, which is short for Fludeoxyglucose-PET [4, 31]). For the majority of contemporary dementia disease diagnosis studies, they mainly rely on conventional pattern recognition tools [20, 32, 35]. For instance, cortical thickness maps are generated from sMRI and Support Vector Machine (SVM) is employed to differentiate *Mild Cognitive Impairment* (MCI) from AD in [32]. In [20], the curse-of-dimension problem commonly seen in pattern recognition studies is investigated in dementia diagnosis, and ensemble classifiers are constructed via sparse encodings for dementia disease prediction. In [35], local volumetric measurements obtained from sMRI are fed into hierarchical networks to discern MCI patients from AD patients. It can be summarized from existing

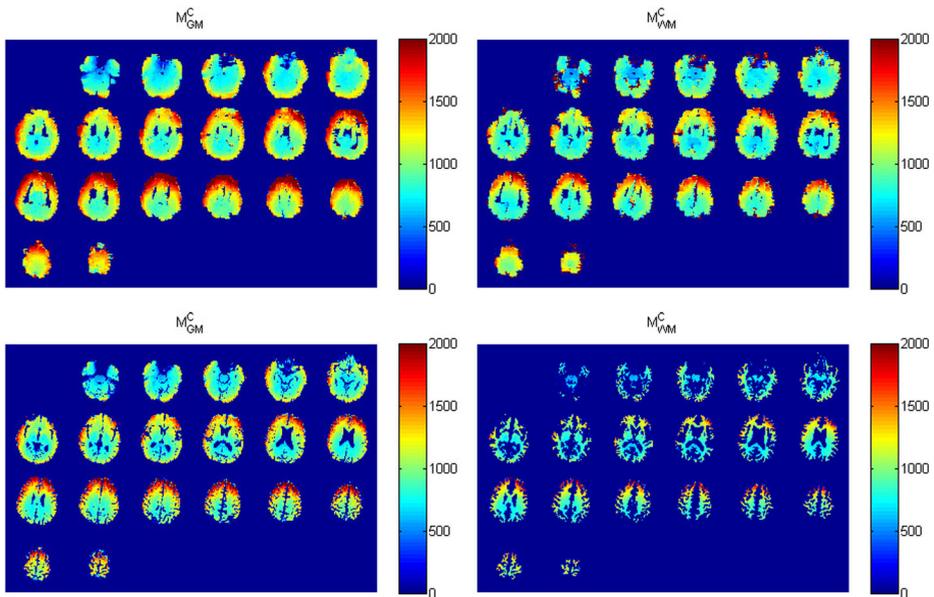


Fig. 2 An Example of PVE correction results via the compared regression-based method (1st row) and the pixel-wise correction method (2nd) on the same patient (the plotting scale unit is [mL/(100g*min)])

studies that, dementia disease prediction is often considered as either a classification or a regression problem.

In this study, a novel dementia diagnosis strategy based on ASL-MRI is introduced for the first time. The strategy is composed of two steps. The first step is to introduce a novel pixel-wise PVE correction method, which only incorporates information extracted from one single pixel for its own PVE correction, rather than information from both a pixel itself and its neighboring pixels commonly adopted in contemporary PVE correction methods [1, 5, 6]. Existing problems such as blurring and brain details loss commonly seen in correction results of those contemporary PVE correction methods can be properly tackled, and under-estimated CBF in ASL-MRI can be well improved. The second step is to present a novel dementia disease prediction method based on corrected ASL-MRI from a new perspective of ranking, instead of the conventional classification and regression viewpoints in contemporary studies [20, 32, 35]. The reason to conduct dementia disease prediction via ranking is also introduced later.

The organization of this paper is as follows. In Section 2, the pixel-wise PVE correction method is elaborated. Then, the dementia disease prediction method via ranking is introduced to fulfill the dementia diagnosis task based on corrected ASL-MRI in Section 3. In Section 4, extensive experiments are conducted to evaluate the performance of the new strategy, with its two critical steps compared with several conventional PVE correction and disease prediction methods. Experimental results of all methods are evaluated from the statistical point of view, and the conclusion of this study is drawn in Section 5. Main contributions of this study can be summarized as: 1) A novel pixel-wise PVE correction method on ASL-MRI, which is capable to tackle problems of blurring and brain details loss in correction results commonly seen in conventional PVE correction methods, is proposed in this paper; 2) The first attempt to diagnose dementia disease based on corrected ASL-MRI from a new ranking perspective is also introduced.

2 A novel pixel-wise PVE correction method on ASL-MRI

The PVE correction problem in ASL-MRI is described as follows. Provided a single pixel i in an ASL-MRI image, its control magnetization M_C and label magnetization M_L can be directly measured from the acquired control and label images, and they can be mathematically represented as:

$$M_C = P_{GM} \cdot M_{GM}^C + P_{WM} \cdot M_{WM}^C + P_{CSF} \cdot M_{CSF}^C \quad (1)$$

$$M_L = P_{GM} \cdot M_{GM}^L + P_{WM} \cdot M_{WM}^L + P_{CSF} \cdot M_{CSF}^L \quad (2)$$

where, P_{GM} , P_{WM} , P_{CSF} indicate the fractional GM, WM, CSF tissue volume on pixel i respectively, and they can be obtained from pre-requisite brain segmentation using the SPM toolbox [29] (i.e., in other words, they are known parameters); M_{\star}^C and M_{\star}^L (i.e., \star represents one of the GM, WM and CSF tissues) denote the control and label magnetization caused by tissue \star on pixel i ; they are unknown parameters to be solved in PVE correction. After PVE correction, ASL signal of tissue \star on the single pixel i can be calculated using $\frac{M_{\star}^C - M_{\star}^L}{M_{\star}^C}$. Thus, CBF of tissue \star on the single pixel i is proportional towards its ASL signal and can be obtained by existing compartment models therein [13].

In order to solve unknowns in (1) & (2), regression techniques have been utilized in [1, 5, 6]. Equation (1) & (2) together construct indefinite equations (i.e., 2 equations, 5

unknowns, given the basic assumption that $M_{CSF}^C = M_{CSF}^L$ as the control and label magnetization of CSF are often equivalent and the number of unknowns in (1) & (2) can be reduced by 1 [10, 17]). Neighbors of pixel i are necessary to be incorporated for adding up extra information for solving the 5 unknowns in regressions. Given an adopted neighbor of size $n \times n$, a regression matrix P of the size $n^2 \times 3$ can be formulated using P_{GM} , P_{WM} , and P_{CSF} , which include fractional GM, WM, and CSF tissue volume of all n^2 neighbor pixels respectively as P 's three columns. Unknowns M_\star^C and M_\star^L on one single pixel can be obtained using $M_\star^C = (P^T P)^{-1} P^T \hat{M}_C$ and $M_\star^L = (P^T P)^{-1} P^T \hat{M}_L$, where \hat{M}_C (\hat{M}_L) depicts a matrix with control (label) magnetization of all n^2 neighbor pixels as its elements; T and -1 represent the transpose and the inverse of a matrix, respectively.

Although the regression process is easy to understand and simple to implement, its shortcomings are obvious. Since pixel neighbors are incorporated in PVE correction, problems of blurring and brain details loss become inevitable in its correction results. A case in point is illustrated in Fig. 2. The 1st row demonstrates solved M_{GM}^C and M_{WM}^C from the transverse view via the regression-based method [1] with a neighbor of size 5×5 . When conducting PVE correction on a single pixel i , its all 25 neighboring pixels are incorporated, and most of them will be utilized again when conducting PVE correction on pixels nearby the single pixel i . Thus, PVE correction results on a single pixel itself together with its neighboring pixels will demonstrate a high degree of similarity, making blurring and brain details loss commonly seen in those correction results. Therefore, imprecisely calculated CBF based on those corrected ASL-MRI cannot help in brain atrophy identification, which is likely to mislead or even deteriorate the following critical dementia disease diagnosis.

In order to tackle the above problem, a novel PVE correction method only incorporating information obtained from one single pixel when correcting its own PVE is presented at this section. Given the basic assumption $M_{CSF}^C = M_{CSF}^L$, (1) & (2) can be firstly re-written as follows:

$$\frac{\Delta M}{M_C} = \frac{M_C - M_L}{M_C} = \frac{P_{GM} \cdot \Delta M_{GM} + P_{WM} \cdot \Delta M_{WM}}{P_{GM} \cdot M_{GM}^C + P_{WM} \cdot M_{WM}^C + P_{CSF} \cdot M_{CSF}^C} \tag{3}$$

where, ΔM represents the difference between the control and label magnetization. Equation (3) is then evaluated using the two following constrained optimization problems:

$$\begin{aligned} \min \sum_{i=1}^N \|M_{C,i} - P_{GM} \cdot M_{GM}^C - P_{WM} \cdot M_{WM}^C - P_{CSF} \cdot M_{CSF}^C\|^2 \\ \text{s.t. } M_{CSF}^C \geq M_{GM}^C \geq M_{WM}^C \end{aligned} \tag{4}$$

$$\begin{aligned} \min \sum_{i=1}^N \|\Delta M_i - P_{GM} \cdot \Delta M_{GM} - P_{WM} \cdot \Delta M_{WM}\|^2 \\ \text{s.t. } \frac{\Delta M_{GM}}{M_{GM}^C} \geq \frac{\Delta M_{WM}}{M_{WM}^C} \end{aligned} \tag{5}$$

where, i denotes the i^{th} ASL-MRI image obtained within a repeated ASL scanning process, which is realized in the clinical scanning protocol of this study to improve the Signal-to-Noise Ratio (SNR) of ASL; constraints in (4) & (5) are based on clinical understandings of brain tissues in ASL [17, 26]. (4) & (5) can be further constructed using Karush-Kuhn-Tucker (KKT) multipliers, and solved following the split-Bregman method [12]. Details of them are elaborated in Table 1.

Table 1 Steps of the Pixel-wise PVE Correction Method

Inputs	P_{GM} P_{WM} P_{CSF} $M_{C,i}$ ΔM_i ($i = 1, \dots, N$)
Initialization	$M_{GM}^{C,0}$ $M_{WM}^{C,0}$ $M_{CSF}^{C,0}$ ΔM_{GM}^0 ΔM_{WM}^0
Steps	
repeat	(at iteration k)
Step 1	$M_{GM}^{C,k} = \min_{M_{GM}^C} \sum_{i=1}^N \ M_{C,i} - P_{GM} \cdot M_{GM}^C - P_{WM} \cdot M_{WM}^{C,k-1} - P_{CSF} \cdot M_{CSF}^{C,k-1}\ ^2 + \lambda_1 \ M_{CSF}^{C,k-1} - M_{GM}^C\ + \lambda_3 \ M_{GM}^C - M_{WM}^{C,k-1}\ $
Step 2	$M_{WM}^{C,k} = \min_{M_{WM}^C} \sum_{i=1}^N \ M_{C,i} - P_{GM} \cdot M_{GM}^{C,k-1} - P_{WM} \cdot M_{WM}^C - P_{CSF} \cdot M_{CSF}^{C,k-1}\ ^2 + \lambda_2 \ M_{CSF}^{C,k-1} - M_{WM}^C\ + \lambda_3 \ M_{GM}^{C,k-1} - M_{WM}^C\ $
Step 3	$M_{CSF}^{C,k} = \min_{M_{CSF}^C} \sum_{i=1}^N \ M_{C,i} - P_{GM} \cdot M_{GM}^{C,k-1} - P_{WM} \cdot M_{WM}^{C,k-1} - P_{CSF} \cdot M_{CSF}^C\ ^2 + \lambda_1 \ M_{CSF}^C - M_{GM}^{C,k-1}\ + \lambda_2 \ M_{CSF}^C - M_{WM}^{C,k-1}\ $
until	$\max(\ M_{GM}^{C,k+1} - M_{GM}^{C,k}\ _\infty, \ M_{WM}^{C,k+1} - M_{WM}^{C,k}\ _\infty, \ M_{CSF}^{C,k+1} - M_{CSF}^{C,k}\ _\infty) \leq tol_1$
Outputs 1	M_{GM}^C M_{WM}^C M_{CSF}^C
repeat	(at iteration l)
Step 4	$\Delta M_{GM}^l = \min_{\Delta M_{GM}} \sum_{i=1}^N \ \Delta M_i - P_{GM} \cdot \Delta M_{GM} - P_{WM} \cdot \Delta M_{WM}^{l-1}\ ^2 + \lambda_4 \left\ \frac{\Delta M_{GM}}{M_{GM}^C} - \frac{\Delta M_{WM}^{l-1}}{M_{WM}^C} \right\ $
Step 5	$\Delta M_{WM}^l = \min_{\Delta M_{WM}} \sum_{i=1}^N \ \Delta M_i - P_{GM} \cdot \Delta M_{GM}^{l-1} - P_{WM} \cdot \Delta M_{WM}\ ^2 + \lambda_4 \left\ \frac{\Delta M_{GM}^{l-1}}{M_{GM}^C} - \frac{\Delta M_{WM}}{M_{WM}^C} \right\ $
until	$\max(\ \Delta M_{GM}^{l+1} - \Delta M_{GM}^l\ _\infty, \ \Delta M_{WM}^{l+1} - \Delta M_{WM}^l\ _\infty) \leq tol_2$
Outputs 2	ΔM_{GM} ΔM_{WM}

Table 1 is composed of detailed steps to correct PVE on one single pixel via this new pixel-wise PVE correction method. In Table 1, Inputs include the fractional GM, WM, CSF tissue volume (i.e. P_{GM} , P_{WM} , and P_{CSF}) on the single pixel to be corrected, the control magnetization of that single pixel residing in the same coordinate of the whole N ASL-MRI images within a repeated scanning process (i.e. $M_{C,i}$, $i = 1, \dots, N$) and the corresponding difference between the control and the label magnetization (i.e. ΔM_i , $i = 1, \dots, N$). For those inputs, the fractional tissues (i.e. P_{GM} , P_{WM} , and P_{CSF}) are obtained from pre-requisite brain segmentation using SPM as mentioned before, and all magnetization measures are obtained directly within the repeated scanning process. Essential steps in Table 1 follows the split-Bregman method, and initialization of unknowns (i.e., M_{GM}^C , M_{WM}^C , M_{CSF}^C , ΔM_{GM} , and ΔM_{WM}) are obtained from correction results of the conventional regression-based method [1] to alleviate unstable influence brought by different settings of initializations in optimizations within the split-Bregman method.

The split-Bregman method in Table 1 is composed of two folds. The first one is to solve for control magnetization of different tissues (i.e., Outputs 1: M_{GM}^C , M_{WM}^C and M_{CSF}^C) via Steps 1 to 3. In each step, one unknown (e.g. M_{GM}^C in Step 1) is separated from the rest via the split-Bregman method, and KKT multiples $\lambda_{(\cdot)}$ are incorporated to form a typical Quadratic Programming (QP) problem. In order to solve the QP problem in each step, an interior point method [34] is incorporated. Iterations from Steps 1 to 3 will terminate when the following stopping criterion meets: $\max(\|M_{GM}^{C,k+1} - M_{GM}^{C,k}\|_\infty, \|M_{WM}^{C,k+1} - M_{WM}^{C,k}\|_\infty, \|M_{CSF}^{C,k+1} - M_{CSF}^{C,k}\|_\infty) \leq tol_1$, in which tol_1 denotes an enough small change between two obtained results within two consecutive iterations k and $k + 1$; $\|\cdot\|_\infty$ represents the spectral norm. After obtaining Outputs 1, similar steps (i.e., Steps 4 to 5) are performed

to calculate ΔM_{GM} and ΔM_{WM} for Outputs 2, following the split-Bregman method as well. After fulfilling all these steps, outcomes obtained from both Outputs 1 and 2 comprise PVE correction results obtained by the new pixel-wise PVE correction method. The 2nd row of Fig. 2 demonstrates corresponding correction results obtained by the new method on the same patient. It can be observed that, problems of blurring and brain details loss which exist in the 1st row obtained by the conventional regression-based method can be tackled well.

Although the new method is more sophisticated than the conventional regression-based method and requires more computational time, parallel computing techniques can be incorporated to improve its efficiency. The merit of the new method in practical implementation is that, PVE correction on each individual pixel only incorporates information obtained from itself, thus PVE corrections on different pixels are totally independent (which is not like the compared conventional method with neighboring pixels incorporated). Therefore, parallel computing can be carried out with the aid of multi-core & multi-thread processors for efficiency boosting in the implementation of the new method. In this study, the average computation time of PVE correction on one patient using the new method with parallel computing techniques is less than 2 mins using an Intel Core i7-3770 CPU (with 4 cores and 8 threads). Thus, the new PVE correction method can effectively handle existing problems of blurring and brain details loss commonly seen in conventional PVE correction methods, as well as efficiently accomplish the PVE correction task within an acceptable response time.

3 A novel dementia disease prediction method on corrected ASL-MRI via ranking

After PVE correction on ASL-MRI is conducted, the next critical step is to predict the dementia disease severity of patients based on their ASL-MRI after PVE correction. Conventional dementia studies often aim to differentiate patients of various dementia disease progressions, including AD, MCI and *Non Cognitive Impairment* (NCI) [4, 20, 31, 32, 35]. Hence, dementia disease prediction in those conventional studies is often considered as either a classification or a regression problem, and popular pattern recognition tools such as SVM and linear/non-linear regressions are often employed [20, 32, 35]. In this section, a novel dementia disease prediction method from a new ranking perspective is presented.

Generally speaking, ranking aims to sort a list of objects according to a system of rating or a record of performance. The intuition to formulate dementia disease prediction via ranking in this study is elaborated as follows. From the conventional classification perspective, such a disease prediction task can be realized by classifying ungraded ASL-MRI into various dementia progression stages (i.e., classes), and their disease severity can be suggested therein [20, 32, 35]. From the conventional regression perspective, such a disease prediction task can be realized within a regression procedure, and disease severities of undiagnosed patients can be revealed by outcomes of the regression process (i.e., often in terms of real numbers). For both classification and regression methods, ASL-MRI images with clinicians' diagnosis results are often utilized in their training processes for tuning unknown parameters in either classifiers or regressors, but none of these images will be explicitly employed in the subsequent yet important phase, the disease prediction process. Therefore, a typical classification or regression procedure does not comply well with conventional clinical decision support techniques (e.g., case-based reasoning), in which new cases are to be diagnosed with reference to other previously diagnosed cases. It is known that, such a case-based reasoning strategy is often relied in real-life clinical diagnosis, and clinicians can improve their decisions or even gain confidence when referring to previously diagnosed cases. Hence, it inspires us to propose a new dementia disease severity prediction method,

which can explicitly incorporate those previously diagnosed cases in the interpolation of new undiagnosed cases, following the introduced conventional clinical decision-making process.

In this study, the disease severity prediction problem is originally considered as a ranking process following the above intuition, and ranking can provide a better fit to the prediction task compared with conventional classification or regression manners. The flowchart of the new “prediction via ranking” method is illustrated in Fig. 3. The main idea is to sort ASL-MRI into a ranked images list, according to the dementia disease severity depicted in all listed ASL-MRI images. The disease severity of a new undiagnosed ASL-MRI can then be interpolated using its neighboring diagnosed ASL-MRI in the ranked list (i.e., Fig. 3). In order to achieve such a ranked images list, a ranking function needs to be determined, so that ASL-MRI images can be sorted into a ranked list based on the ranking function. In this section, a novel learning method inspired by a conventional position-based ranking evaluation measure, the Normalized Discounted Cumulative Gain (NDCG) [15], is introduced to fulfill the ranking function learning task.

The explicit form of NDCG is generally described as below:

$$NDCG = N_M^{-1} \times DCG = N_M^{-1} \sum_{x \in \chi} \frac{2^{r(x)} - 1}{\log_2(1 + \pi(x))} \tag{6}$$

where, x is an ASL-MRI image and χ denotes the set of ASL-MRI images to be ranked; $r(x)$ and $\pi(x)$ are the annotated dementia disease severity by clinicians (i.e. often in the form of integer grades) and the position of image x within the ranked image list, respectively; N_M is a normalization term denoting the maximum of DCG, which can be obtained when all images are sorted in a perfect order of decreasing severity of dementia disease. Therefore, the range of NDCG is within $[0, 1]$, where the lower bound and upper bound denote a perfect increasing severity order and a perfect decreasing severity order, respectively.

In order to learn a ranking function based on (6), an optimization process needs to be conducted. Unfortunately, optimization cannot be directly applied on (6) for learning ranking

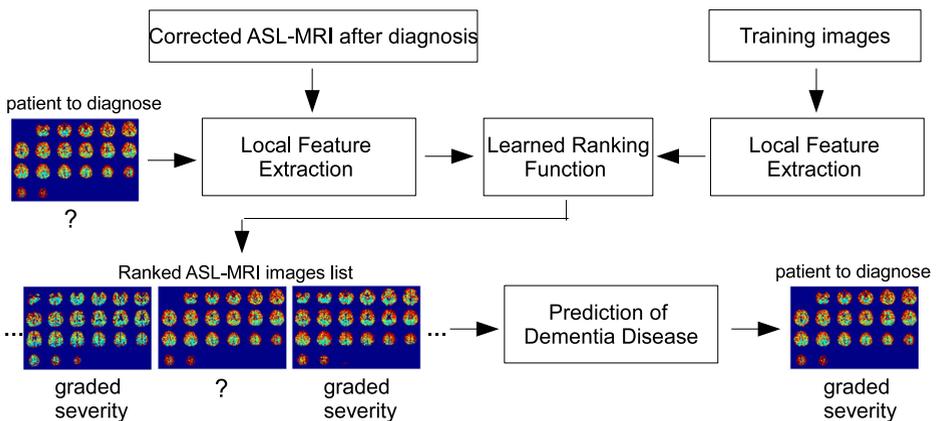


Fig. 3 Flowchart of the “prediction via ranking” method

functions, as NDCG itself is neither continuous nor differentiable in terms of the discrete position term $\pi(x)$. Thus, position $\pi(x)$ needs to be revised and is first approximated as follows:

$$\pi(x) \simeq 1 + \sum_{y \neq x, y \in \chi} \text{sign}(s_y - s_x) = 1 + \sum_{y \neq x, y \in \chi} \text{sign}(f(\hat{y}) - f(\hat{x})) \tag{7}$$

where, \hat{x} represents a d -dimensional feature vector extracted from the ASL-MRI image x ; s_x is the score of image x calculated from the ranking function $f(\hat{x})$, which is of a linear form in this study (i.e. $f(\hat{x}) = \langle \theta, \hat{x} \rangle$, where \langle, \rangle denotes an inner product between θ and \hat{x}). Hence, θ is also a d -dimensional vector and there are d unknown parameters in it to be learned). $\text{sign}(s_y - s_x)$ is an signum function, whose value is positive when $s_y \geq s_x$ and negative otherwise. The reason to revise position $\pi(x)$ as (7) is as follows. When the score of image x is smaller than that of image y (i.e. $s_x < s_y$), $\text{sign}(s_y - s_x)$ becomes positive and $\pi(x)$ becomes larger due to (7), which matches the fact that images reflecting lighter disease severity (i.e., indicated by smaller score s_x) should be ranked in the rear of a ranked image list (i.e. indicated by a larger value of position $\pi(x)$), in a demanded descending order of disease severity.

Moreover, step transition characteristics of the signum function in (7) makes direction optimization on (6) still infeasible to be implemented for learning the ranking function $f(\hat{x})$. Thus, the signum function in (7) can be further approximated using $\text{sign}(\zeta) \simeq \frac{\zeta}{\sqrt{\zeta^2 + \alpha^2}}$, where ζ denotes the variable of the signum function, and α controls the sharpness of the approximated function towards $\text{sign}(\zeta)$. It is shown in Fig. 4 that, the less α becomes,

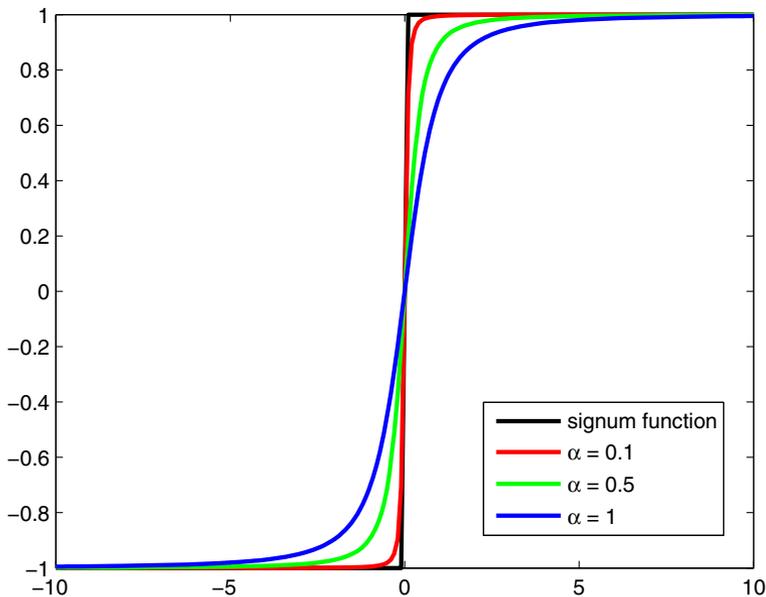


Fig. 4 An illustration of approximating signum function with different α

the more similar it turns towards the signum function. In this way, a new continuous approximated position $\pi'(x)$ can be described as:

$$\begin{aligned} \pi'(x) &\simeq 1 + \sum_{y \neq x, y \in \chi} \text{sign}(s_y - s_x) = 1 + \sum_{y \neq x, y \in \chi} \text{sign}(s_{yx}) = 1 \\ &+ \sum_{y \neq x, y \in \chi} \frac{s_{yx}}{\sqrt{s_{yx}^2 + \alpha^2}}, \quad s_{yx} = s_y - s_x \end{aligned} \tag{8}$$

In this way, a new continuous and differentiable approximation towards NDCG (i.e., denoted as ‘‘C-NDCG’’ in this study) can be explicitly depicted as:

$$\text{C-NDCG}(x) = N_M^{-1} \sum_{x \in \chi} \frac{2^{r(x)} - 1}{\log_2 \left(2 + \sum_{y \neq x, y \in \chi} \frac{s_{yx}}{\sqrt{s_{yx}^2 + \alpha^2}} \right)} \tag{9}$$

A corresponding algorithm to directly optimize C-NDCG via a gradient method for ranking functions learning is listed in Table 2. The critical step here is to calculate the gradient of C-NDCG with respect to the learned parameter θ (i.e., $\frac{\partial \text{C-NDCG}(x)}{\partial \theta}$) in Steps T4&T5 of

Table 2 Steps of ranking functions learning on C-NDCG

Inputs	<ol style="list-style-type: none"> 1. ASL-MRI images for training: $\{x \in \chi\}$ 2. ASL-MRI images for validation: $\{x_v \in \chi_v\}$ 3. Number of Iterations: T 4. Learning rate: η
Training	<ol style="list-style-type: none"> T1. Initialize parameter θ of the ranking function $f(\hat{x})$ as θ_0 T2. For t = 1 to T T3. Set $\theta = \theta_{t-1}$ T4. Feed $\{x \in \chi\}$ to (10) to calculate the gradient T5. Update θ via gradient ascent: $\theta = \theta + \eta \cdot \frac{\partial \text{C-NDCG}(x)}{\partial \theta}$ T6. Set $\theta_t = \theta$ T7. End for T2
Training Validation	<ol style="list-style-type: none"> T. T learned ranking functions $f(\hat{x})$ with T corresponding learned parameters θ V1. For j = 1 to T V2. Feed j^{th} learned ranking function $f_j(\hat{x})$ to $\{x_v \in \chi_v\}$ to rank validation images V3. Calculate its corresponding NDCG value using (6) V4. End for V1 V5. Determine $f_{opt}(\hat{x})$ as the one with the highest NDCG value
Outputs	Optimal learned ranking function: $f_{opt}(\hat{x})$

Table 2. Detailed derivation is elaborated in the Appendix of the paper. The gradient can be computed as:

$$\frac{\partial \text{C-NDCG}(x)}{\partial \theta} = N_M^{-1} \sum_{x \in \mathcal{X}} \left(-\frac{2^{r(x)} - 1}{\log_2^2(1 + \pi'(x))} \cdot \frac{1}{(1 + \pi'(x)) \ln 2} \right) \times \left(\sum_{y \neq x, y \in \mathcal{X}} \frac{\alpha^2}{(s_{yx}^2 + \alpha^2)^{\frac{3}{2}}} \cdot \left(\frac{\partial f(\hat{y})}{\partial \theta} - \frac{\partial f(\hat{x})}{\partial \theta} \right) \right) \tag{10}$$

Since the local optimizer of a gradient method cannot always guarantee the global optimal solution, we run T iterations of ranking functions learning θ_t initialized by previously learned $\theta_{(t-1)}$, where t denotes the t -th iteration. Hence, after conducting the training phase in Table 2, there are T ranking functions learned with their corresponding learned θ . Then, a validation phase is incorporated afterwards to select an optimal ranking function $f_{opt}(\hat{x})$ from those T candidates, as the one with the highest NDCG value after applying all T learned ranking functions obtained from the training phase to rank the validation set of images (i.e., calculated using (6)).

In the testing phase, an undiagnosed ASL-MRI image x is then sorted together with other diagnosed images with clinicians’ annotated grades indicating their disease severities into a ranked image list using $f_{opt}(\hat{x})$. Hence, graded information (of those diagnosed cases) is explicitly utilized, which mimics the conventional case-based reasoning procedure. Grade g_{x_i} of the ASL-MRI image x located at position i of the ranked images list can then be interpolated using both calculated scores from itself (s_{x_i}) and its neighboring images ($s_{x_{i-1}}, s_{x_{i+1}}$) as well as their annotated grades ($g_{x_{i-1}}$ and $g_{x_{i+1}}$), which are known diagnosis results to clinicians. The grade g_{x_i} of an undiagnosed ASL-MRI image x can then be explicitly described using the following piecewise function:

$$g_{x_i} = \begin{cases} g_{x_{i+1}} & \text{if } g_{x_{i+1}} = g_{x_{i-1}} \\ g_{x_{i+1}} + \frac{s_{x_i} - s_{x_{i+1}}}{s_{x_{i-1}} - s_{x_{i+1}}} \times (g_{x_{i-1}} - g_{x_{i+1}}) & \text{if } g_{x_{i-1}} > g_{x_{i+1}} \\ g_{x_{i-1}} + \frac{s_{x_{i-1}} - s_{x_i}}{s_{x_{i-1}} - s_{x_{i+1}}} \times (g_{x_{i+1}} - g_{x_{i-1}}) & \text{if } g_{x_{i-1}} < g_{x_{i+1}} \end{cases} \tag{11}$$

where, $s_{x_i} = f_{opt}(\hat{x}_i)$, $s_{x_{i-1}} = f_{opt}(\hat{x}_{i-1})$, and $s_{x_{i+1}} = f_{opt}(\hat{x}_{i+1})$. It can be easily perceived that, when two neighboring ranked images are of the same disease severity (i.e., $g_{x_{i+1}} = g_{x_{i-1}}$), the undiagnosed image should share the same severity as them in the ranked images list; when two neighboring ranked images are of different disease severities, the severity of the undiagnosed image is to be determined by both results of the learned ranking function (i.e., scores) and previously annotated information (i.e., grades), which complies well with the conventional case-based reasoning procedure.

4 Experiments and analysis

4.1 Data description and pre-processings

In order to demonstrate the superiority of the newly proposed dementia diagnosis strategy, clinical data obtained from 350 real patients, including 110 AD patients, 120 MCI patients and 120 NCI patients (as normal controls) acquired in the affiliated hospital of Nanchang University, is utilized. Informed consent is obtained from all patients for research purpose. The averaged age of these patients is 70.56 ± 7.20 years old. In ASL-MRI images

acquisition, A SIEMENS 3T TIM Trio MR scanner is utilized and 23 ASL-MRI images are acquired consecutively for each patient to improve the SNR of ASL within a repeated scanning process (i.e., $N = 23$ in (4) & (5) and relevant Eqs in Table 1). Other acquisition parameters include: labeling duration = 1500 ms, post-labeling delay = 1500ms, TR/TE = 4000/9.1ms, ASL voxel size = $3 \times 3 \times 5\text{mm}^3$. In PVE correction, pre-defined parameters include: $\lambda_1 = \lambda_2 = \lambda_3 = 0.1$, $\lambda_4 = 0.01$, $tol_1 = 10$, $tol_2 = 0.5$, the adopted neighbor size is 9×9 when implementing the regression-based PVE correction method for initializations in Table 1. High-resolution MPRAGE (i.e., which is short for Magnetization Prepared Rapid Acquisition Gradient Echo) T1-weighted MRI images [2] are also acquired for all patients simultaneously in their scannings. After MPRAGE and ASL-MRI images acquisition, brain extraction and motion correction are applied on acquired MPRAGE and ASL-MRI images. The MPRAGE image of every patient is then segmented into GM/WM/CSF components with their probability maps P_{GM} , P_{WM} , and P_{CSF} generated for inputs in Table 1, using the SPM toolbox [29]. The above obtained maps are then co-registered towards their corresponding ASL-MRI images after motion correction for every patient using the FSL toolbox [7].

Experiments in this study are divided into two aspects. The first is to incorporate all patients' ASL-MRI images to demonstrate the superiority of the new pixel-wise PVE correction method (i.e., Step 1 of the newly introduced strategy), in comparison with other conventional PVE correction methods (Section 4.2). The second is to apply the new "prediction via ranking" method on ASL-MRI after PVE correction to diagnose dementia disease (i.e., Step 2 of the newly introduced strategy), in comparison with several popular pattern recognition tools to reveal the superiority of the new ranking method (Section 4.3). Comprehensive statistical analysis is performed in all experiments.

4.2 Experiments and analysis on PVE correction

As introduced in Section 1, ASL-MRI images suffering from PVE will often result in under-estimated CBF, and superior PVE correction methods will often improve CBF after PVE correction better. The newly proposed pixel-wise PVE correction method (denoted as "New") is compared with the conventional regression-based PVE correction method (denoted as "RB"), regarding calculated CBF from their corrected ASL-MRI results. For compared RB methods, different sizes of neighbors are implemented in experiments, including sizes of 9×9 (i.e., to illustrate a small size of neighborhood), and 15×15 (i.e., to illustrate a large size of neighborhood). They are denoted as "RB-9" and "RB-15", respectively.

Figure 5 displays PVE correction results of an AD patient obtained by all compared PVE correction methods. In this figure, each row represents correction results obtained by one method. For columns of Fig. 5, the first column depicts the solved control magnetization in GM (i.e. M_{GM}^C), the second column displays the solved control magnetization in WM (i.e. M_{WM}^C), the third column demonstrates the solved control magnetization in CSF (i.e. M_{CSF}^C), and the fourth column illustrates histograms of subject-wise averaged ASL signal $\frac{\Delta M_{GM}}{M_{GM}^C}$ (red) and $\frac{\Delta M_{WM}}{M_{WM}^C}$ (blue), as well as the calculated voxel-wise ratio of GM flow towards WM flow annotated on tops of histograms. It can be easily observed that, blurring and brain details loss in correction results of "New" can be greatly improved. Also, the ratio of GM flow to WM flow obtained by "New" is 2.0314, which complies well with clinical literatures [17, 26] compared with other methods (e.g., ratio results of "RB-9" and "RB-15" are 1.6853, and 1.7601, respectively in Fig. 5, which are significantly smaller and do not

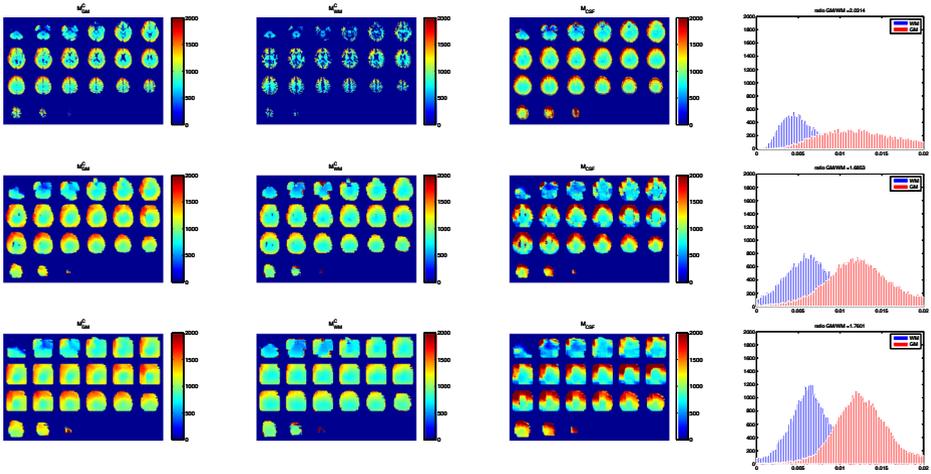


Fig. 5 PVE correction results on an AD patient obtained by different methods (1st row: New; 2nd row: RB-9; 3rd row: RB-15)

comply well with clinical literatures). Similar results can be observed in all 350 patients as well. After calculating CBF of all patients based on correction results obtained from all PVE correction methods, a box-and-whisker plot summarizing CBF of all patients is illustrated in Fig. 6. In each box, a horizontal line is drawn across the box at the median of CBF, while the upper- and lower-quartiles of CBF are depicted as lines above and below the median. A vertical dashed line is drawn up from the upper-quartile and down from the lower-quartile to their most extreme data points, which are within a distance of 1.5 Inter-Quartile Range (IQR) [28]. Each data point beyond ends of the vertical line is marked via a plus sign. It can

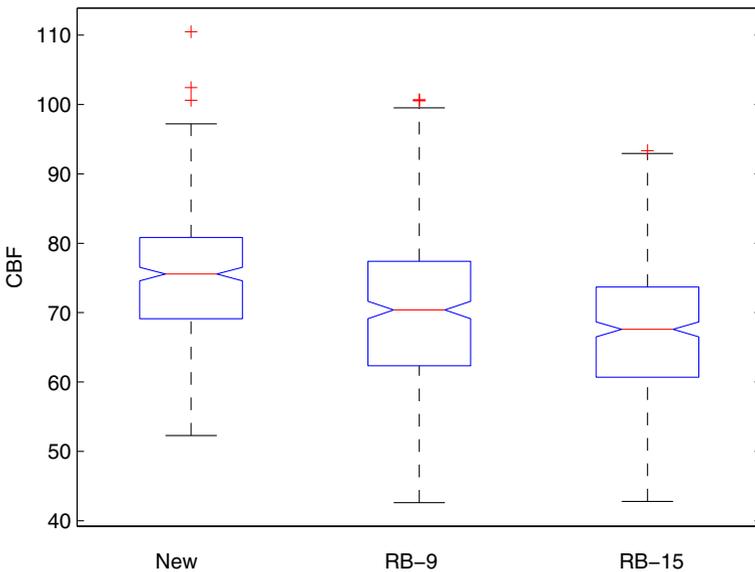


Fig. 6 Box plot of CBF obtained by all compared PVE correction methods

be observed that, the box of “New” is significantly higher than those of others, which indicates that the under-estimated CBF in original ASL-MRI images can be improved better by the new pixel-wise correction method.

In order to reveal the superiority of “New” in PVE correction from the statistical point of view, a statistical test made up of one-way Analysis Of Variance (ANOVA) followed by post-hoc multiple comparison tests [28] is further utilized for statistical analysis on all CBF results obtained by all compared PVE correction methods on all patients data. ANOVA is a popular correction of models analyzing the difference between diverse group means and their associated variations in statistics [28]. In one-way ANOVA, CBF results obtained from all methods are compared to test a hypothesis (H_0) that “CBF means of various methods are equivalent”, against the general alternative that these means cannot be all the same. P-value is used here as an indicator to reveal whether H_0 holds or not. In this study, p-values calculated from all CBF results are nearly 0, which is a strong indication that all these methods cannot share the same CBF means. Therefore, the next step is to conduct more detailed paired comparisons. The reason to do so is because that, the alternative against H_0 is too general. Information about which method is superior from the statistical perspective cannot be perceived by one-way ANOVA alone. There are two kinds of evaluation after applying multiple comparison tests on calculated CBF of all methods, and quantitative evaluation results are shown in Table 3. For the two kinds of evaluation, one is estimated CBF mean difference, which is a single-value estimator of CBF mean difference. Another is a 95 % Confidence Interval (CI). In statistics, a CI is a special form of interval estimator for a parameter (i.e. CBF mean difference in this experiment). Generally speaking, instead of estimating the parameter by a single value, CI is capable to provide an interval estimation which is likely to include the estimated parameter within a specified interval. To be specific, “New” is 4.3904 higher than “RB-9”. The CBF mean difference (i.e., using “New” minus “RB-9”) is likely to fall within a 95 % CI [2.2873, 6.4935]. Since the upper and lower bounds of the CI are both positive, it gives a strong indication (> 95 %) that the CBF mean difference should be positive. Thus, “New” is superior to “RB-9” in terms of CBF from statistical point of view. For comparisons between “New” and “RB-15”, a similar conclusion can be drawn from Table 3. To sum up, based on the above statistical analysis, the pixel-wise PVE correction method in the newly introduced dementia diagnosis strategy outperforms compared conventional regression-based methods, from the statistical perspective.

4.3 Experiments and analysis of disease severity prediction

After PVE correction has been fulfilled, the next critical step is to conduct dementia disease severity prediction. In this section, the “prediction via ranking” method (denoted as “Ranking”) is evaluated for its capability of dementia disease prediction, based on ASL-MRI after PVE correction. Its pre-defined parameters are set as $T = 200$ and $\eta = 0.01$ in

Table 3 Multiple comparison test of obtained CBF based on Corrected ASL-MRI from all PVE correction methods

Method I	Method II	CBF Mean Difference (I-II)	a 95 % Confidence Interval
New	RB-9	4.3904	[2.2873, 6.4935]
New	RB-15	2.5758	[0.4727, 4.6789]
RB-9	RB-15	-1.8146	[-3.9177, 0.2885]

Table 2 through trial-and-error for optimal performance; α in (8) controlling the sharpness of the approximated function towards the signum function is set as 0.1. Mean ASL signal calculated from the segmented left & right hippocampus, the left & right parahippocampal gyrus, the left & right putamen, and the left & right thalamus (i.e. the above tissue segmentation is realized via IBA-SPM [14]) is utilized to construct a 8-dimensional feature vector \hat{x} in (7) for each patient. Popular pattern recognition tools widely utilized in conventional dementia diagnosis studies, including support vector machine (denoted as “SVM” as a classification tool), support vector regression (denoted as “SVR” as a non-linear regression tool) and linear regression (denoted as “LR”), are implemented in this experiment based on ASL-MRI images of all patients after PVE correction, for dementia diagnosis. Since learning is incorporated in “Ranking”, parameters of other methods are also learned for the sake of fairness. For “SVM” and “SVR”, Gaussian Radial Basis Function (RBF) are adopted as kernels; Gaussian widths are learned via the popular radius/margin bound algorithm [18], and SVM-light toolbox [16] is utilized for their implementations. For “LR” and “Ranking”, disease severities of AD, MCI and NCI are labeled as 1, 2 and 3 respectively. Regression coefficients in “LR” are determined via labels and regressors (i.e., the 8-dimensional feature vector) of the training data.

The whole dataset of 350 patients is equally divided into 5 subsets to conduct a 5-fold cross validation for statistical evaluation. In each subset, patients with different dementia disease severities are roughly equivalent (i.e., 22 AD/ 24 MCI/ 24 NCI in each subset). Since there are training, validation and testing phases in the “Ranking” method and the number of subsets utilized in them are 3, 1 and 1 individually in each trial of the 5-fold cross validation, the total number of trials in the whole 5-fold cross validation is $C_5^3 \cdot C_1^1 \cdot C_1^1 = 20$, where $C_{(*)}^{(*)}$ denotes the number of combinations of \star objects from a set of \star objects. For other compared methods without the validation phase (i.e., “SVM”, “SVR” and “LR”), all non-testing subsets (i.e., training+validation subsets) are utilized for parameters learning in each trial. In order to perceive comprehensive understandings of the newly proposed dementia diagnosis strategy, all compared disease severity prediction methods are tested on all corrected ASL-MRI obtained by all compared PVE correction methods in Sections 4.2, and the combination of PVE correction and disease prediction methods producing the best performance in dementia diagnosis can be suggested therein.

Statistical results of dementia disease prediction obtained by all combinations of PVE correction methods and disease severity prediction methods are elaborated in Table 4. Each entry in Table 4 contains the mean±standard deviation of the difference between predicted disease severity (i.e., grade) and its ground truth annotated by our senior clinicians determined by consensus, as well as the prediction accuracy (i.e., in percentage) in which one predicted case is considered to be accurate if the difference between its prediction and ground truth is less than 0.3 (i.e., suggested by our senior clinicians based on clinical evidence). The least prediction error and the highest prediction accuracy in each trial are highlighted. It can be observed from Table 4 that, the combination of the pixel-wise PVE correction method (“New”) and the “prediction via ranking” method (“Ranking”) achieves the highest prediction accuracy and the least prediction error in most trials (i.e. 17 out of 20 trials with the highest prediction accuracy, and 15 out of 20 trials with the least prediction error). We investigated trials that the newly proposed ranking method cannot perform the best, and found that characteristics between different patients vary more in those trials. Since ranking is more generative than conventional classification, it is more convenient for classification models to tackle those more discriminant cases. However, based on all statistics, the generative ranking method still achieves the most satisfactory outcomes. To

Table 4 Statistical results of dementia disease prediction by all compared methods on ASL-MRI of all patients after PVE correction

Trial	Ranking				SVM			
	New	RB-9	RB-15	New	RB-9	RB-15	RB-9	RB-15
1	0.0317±0.0186 100 %	0.0830±0.0495 92.86 %	0.0790±0.0458 95.71 %	0.1886±0.1128 81.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %
2	0.0660±0.0401 97.14 %	0.0795±0.0442 95.71 %	0.0792±0.0465 92.86 %	0.2810±0.1217 52.86 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %
3	0.1382±0.1079 88.57 %	0.0848±0.0463 92.86 %	0.0710±0.0422 98.57 %	0.1886±0.1128 81.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %
4	0.0737±0.0495 95.71 %	0.1072±0.0628 78.57 %	0.1923±0.1188 57.14 %	0.2429±0.1070 48.57 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %
5	0.1226±0.0689 88.57 %	0.0745±0.0428 98.57 %	0.0713±0.0439 95.71 %	0.2810±0.1217 52.86 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %
6	0.0464±0.0303 98.57 %	0.0930±0.0595 88.57 %	0.1023±0.0569 81.43 %	0.2429±0.1070 48.57 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %
7	0.0441±0.0262 100 %	0.0788±0.0456 97.14 %	0.0738±0.0448 97.14 %	0.1886±0.1128 81.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %	0.1657±0.0939 64.29 %	0.1653±0.0928 61.43 %
8	0.1242±0.1240 88.57 %	0.0942±0.0523 88.57 %	0.0934±0.0523 85.71 %	0.3105±0.1251 41.43 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %
9	0.0838±0.0397 100 %	0.0714±0.0394 100 %	0.0796±0.0463 92.86 %	0.2810±0.1217 52.86 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %	0.2182±0.1190 38.57 %	0.2275±0.1166 40.00 %
10	0.0653±0.0327 100 %	0.0671±0.0406 98.57 %	0.0943±0.0573 88.57 %	0.3105±0.1251 41.43 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %
11	0.0809±0.0526 95.71 %	0.0873±0.0507 94.29 %	0.1010±0.0568 84.29 %	0.2429±0.1070 52.86 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %	0.2142±0.1035 40.00 %	0.2133±0.1036 40.00 %
12	0.0510±0.0337 98.57 %	0.0802±0.0445 98.57 %	0.1048±0.0586 82.86 %	0.3105±0.1251 41.43 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %	0.2941±0.1103 20.00 %	0.3151±0.1153 18.57 %
13	0.0237±0.0116 100 %	0.0880±0.0517 91.43 %	0.0793±0.0470 94.29 %	0.0896±0.0543 48.57 %	0.1886±0.1128 50.00 %	0.1886±0.1128 50.00 %	0.1886±0.1128 50.00 %	0.1886±0.1128 50.00 %
14	0.0447±0.0264 100 %	0.1027±0.0555 82.86 %	0.0816±0.0477 91.43 %	0.1190±0.0656 98.57 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %
15	0.1565±0.0896 84.29 %	0.0819±0.0446 97.14 %	0.0748±0.0451 94.29 %	0.1312±0.0755 97.14 %	0.2810±0.1217 25.71 %	0.2810±0.1217 25.71 %	0.2810±0.1217 25.71 %	0.2810±0.1217 25.71 %
16	0.0626±0.0313 100 %	0.0887±0.0513 95.71 %	0.0796±0.0451 97.14 %	0.1190±0.0656 87.14 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %
17	0.1035±0.0702 92.86 %	0.1107±0.0682 81.43 %	0.1237±0.0660 68.57 %	0.1243±0.0731 91.43 %	0.2429±0.1070 35.71 %	0.2429±0.1070 35.71 %	0.2429±0.1070 35.71 %	0.2429±0.1070 35.71 %
18	0.0620±0.0301 98.57 %	0.0963±0.0540 82.86 %	0.0866±0.0503 92.86 %	0.1190±0.0656 97.14 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %	0.2641±0.1176 30.00 %
19	0.0478±0.0269 100 %	0.0899±0.0524 90.00 %	0.0826±0.0513 94.29 %	0.1593±0.0905 75.71 %	0.3033±0.1151 22.86 %	0.3068±0.1151 17.14 %	0.3033±0.1151 22.86 %	0.3068±0.1151 17.14 %
20	0.0659±0.0376 98.57 %	0.1010±0.0674 87.14 %	0.0858±0.0493 90.00 %	0.1611±0.0780 84.29 %	0.2699±0.1130 21.43 %	0.2766±0.1218 27.14 %	0.2699±0.1130 21.43 %	0.2766±0.1218 27.14 %
Average	0.0747±0.0474 96.29 %	0.0880±0.0712 91.64 %	0.0918±0.0736 88.79 %	0.1546±0.0684 70.07 %	0.1877±0.0801 36.71 %	0.1926±0.0808 36.29 %	0.1877±0.0801 36.71 %	0.1926±0.0808 36.29 %

Table 4 (continued)

Trial	SVR				LR							
	New	RB-9	RB-15	New	RB-9	RB-15	New	RB-9	RB-15			
1	0.1380±0.0783	91.43 %	0.1848±0.1078	55.71 %	0.1698±0.1125	64.29 %	0.2999±0.2237	32.86 %	0.3619±0.2610	31.43 %	0.3291±0.2399	31.43 %
2	0.1916±0.0969	75.71 %	0.2120±0.1018	35.71 %	0.2637±0.1207	31.43 %	0.2205±0.2481	37.14 %	0.1458±0.2390	18.57 %	0.1770±0.2436	20.00 %
3	0.1723±0.1081	77.14 %	0.1991±0.1208	52.86 %	0.2081±0.1345	57.14 %	0.2999±0.2237	32.86 %	0.3619±0.0939	31.43 %	0.3291±0.2399	31.43 %
4	0.2194±0.1141	68.57 %	0.2951±0.1298	30.00 %	0.2647±0.1298	37.14 %	0.1770±0.0742	58.57 %	0.1351±0.2579	31.43 %	0.1254±0.2036	37.14 %
5	0.1957±0.0979	74.29 %	0.2077±0.1031	38.57 %	0.2529±0.1247	37.14 %	0.2205±0.2481	37.14 %	0.1458±0.2390	18.57 %	0.1770±0.2436	20.00 %
6	0.2248±0.1137	65.71 %	0.2524±0.1303	41.43 %	0.2871±0.1393	35.71 %	0.1770±0.1742	57.14 %	0.1351±0.2579	31.43 %	0.1254±0.2036	37.14 %
7	0.1387±0.1125	90.00 %	0.1653±0.1079	62.86 %	0.2012±0.1291	51.43 %	0.2999±0.2237	32.86 %	0.3619±0.2610	31.43 %	0.3291±0.2399	31.43 %
8	0.1716±0.0934	75.71 %	0.3048±0.1258	25.71 %	0.2620±0.1203	31.43 %	0.1470±0.1881	60.00 %	0.1393±0.1713	51.43 %	0.1108±0.1595	51.43 %
9	0.1987±0.0998	74.29 %	0.2143±0.1082	42.86 %	0.2570±0.1353	40.00 %	0.2205±0.2481	37.14 %	0.1458±0.2390	18.57 %	0.1770±0.2436	20.00 %
10	0.1686±0.0919	77.14 %	0.2904±0.1290	31.43 %	0.2782±0.1190	28.57 %	0.1470±0.1881	60.00 %	0.1393±0.1713	51.43 %	0.1108±0.1595	51.43 %
11	0.2256±0.1152	65.71 %	0.3038±0.1332	28.57 %	0.2686±0.1327	35.71 %	0.1770±0.1742	57.14 %	0.1351±0.2579	31.43 %	0.1254±0.2036	37.14 %
12	0.1684±0.0922	77.14 %	0.3050±0.1255	25.71 %	0.2830±0.1151	24.29 %	0.1470±0.1881	60.00 %	0.1393±0.1713	51.43 %	0.1108±0.1595	51.43 %
13	0.1997±0.1223	75.71 %	0.1883±0.1101	52.86 %	0.1686±0.1051	58.57 %	0.2999±0.2237	32.86 %	0.3619±0.2610	31.43 %	0.3291±0.2399	31.43 %
14	0.2384±0.1111	64.29 %	0.2368±0.1093	31.43 %	0.2579±0.1097	30.00 %	0.2044±0.2527	45.71 %	0.1024±0.1991	48.57 %	0.1341±0.1815	52.86 %
15	0.1975±0.1002	74.29 %	0.2327±0.1053	32.86 %	0.2656±0.1182	30.00 %	0.2205±0.2481	37.14 %	0.1458±0.2390	28.57 %	0.1770±0.2436	30.00 %
16	0.2273±0.1125	68.57 %	0.2304±0.1084	32.86 %	0.2572±0.1093	28.57 %	0.2044±0.2527	45.71 %	0.1024±0.1991	48.57 %	0.1341±0.1815	52.86 %
17	0.2196±0.1121	67.14 %	0.2766±0.1320	34.29 %	0.2983±0.1318	24.29 %	0.1770±0.1742	47.14 %	0.1351±0.2579	31.43 %	0.1254±0.2036	37.14 %
18	0.2273±0.1123	68.57 %	0.2394±0.1092	27.14 %	0.2685±0.1087	24.29 %	0.2044±0.2527	45.71 %	0.1024±0.1991	48.57 %	0.2641±0.1815	52.86 %
19	0.1679±0.0921	77.14 %	0.2984±0.1278	30.00 %	0.2583±0.1153	28.57 %	0.1470±0.1881	60.00 %	0.1393±0.1713	51.43 %	0.1108±0.1595	51.43 %
20	0.2305±0.1116	67.14 %	0.2388±0.1099	31.43 %	0.2593±0.1105	28.57 %	0.2044±0.2527	45.71 %	0.1024±0.1991	48.57 %	0.1341±0.1815	52.86 %
Average	0.1461±0.0744	73.79 %	0.1938±0.0868	37.21 %	0.2015±0.0911	36.36 %	0.1698±0.1174	56.57 %	0.2369±0.1257	26.29 %	0.2110±0.1056	28.57 %

be specific, the average prediction accuracy of the newly introduced dementia disease diagnosis strategy in this study (i.e., “New”+“Rank”) is 96.29 %, and 0.0747 ± 0.0474 as its mean and standard deviation of prediction errors. The above outcomes are superior to ones of other combinations based on all patients data.

Another interesting thing to notice is that, for one specific prediction method (e.g. “Ranking”), corrected ASL-MRI obtained by the pixel-wise correction method (e.g., “New”) can help to provide better dementia disease prediction performance than ones obtained by compared conventional regression-based methods (i.e., “RB-9” and “RB-15”). For instance, the averaged prediction accuracy is 96.29 % for “New”+“Ranking”, compared with 91.64 % for “RB-9”+“Ranking” and 88.79 % for “RB-15”+“Ranking”. Similar conclusions can also be drawn in each of other disease severity prediction methods (e.g., “SVM”, “SVR” and “LR”) when comparing prediction results based on different PVE correction methods. All these observations demonstrate the effectiveness of the pixel-wise PVE correction method in differentiating patients with various dementia disease severities in both conventional prediction methods and the newly introduced “prediction via ranking” method. Also, for corrected ASL-MRI obtained by one specific PVE correction method, the newly introduced “prediction via ranking” method outperforms other conventional prediction methods. For instance, based on corrected ASL-MRI obtained by “New”, the dementia diagnosis prediction accuracy of “New”+“Ranking”, “New”+“SVM”, “New”+“SVR” and “New”+“LR” are 96.29 %, 70.07 %, 73.79 % and 36.57 %, respectively. Thus, the effectiveness of the newly introduced “prediction via ranking” method can also be verified.

In Fig. 7, a histogram depicting the distribution of prediction errors obtained by the newly presented dementia disease diagnosis strategy in this study (i.e., “New”+“Rank”) is illustrated based on all diagnosis results obtained from the 5-fold cross validation on all patients data. The number of testing data in Fig. 7 is 1400 (which equals to 350×4 , as each testing subset will be utilized $C_4^3 \times C_1^1 = 4$ times brought by different combinations of training

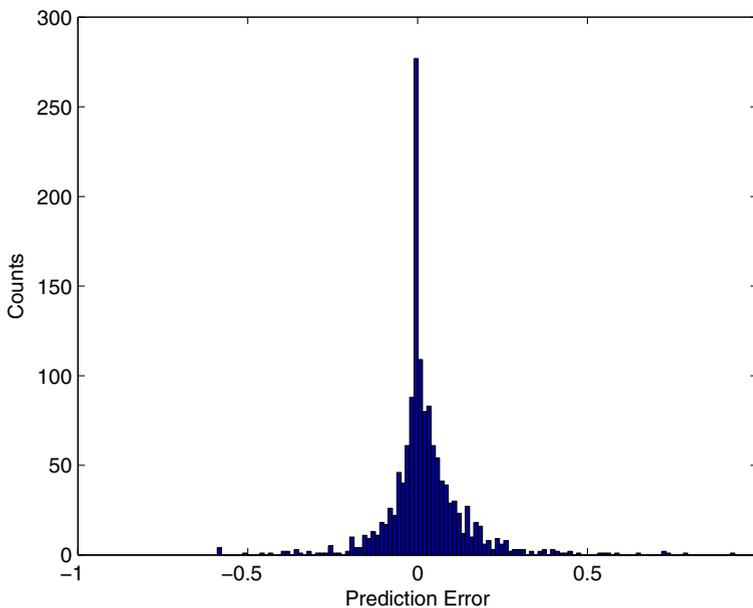


Fig. 7 Histogram of prediction errors obtained by the newly proposed dementia disease diagnosis strategy

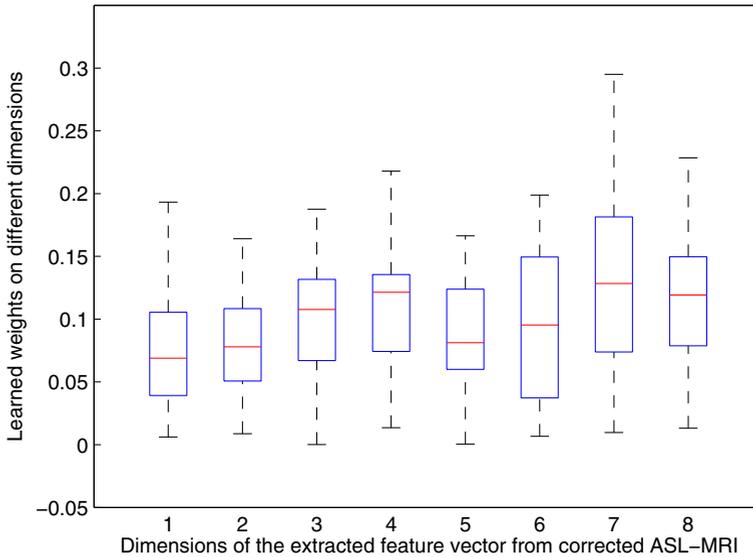


Fig. 8 Summary of learned weights on different dimensions of the extracted feature according to outcomes in “New”+“Rank”

and validation subsets in the 5-fold cross validation). It can be observed that, the prediction error of most cases is within the range $[-0.5, 0.5]$. In Fig. 8, the distribution of all 8 elements of the learned θ in the utilized linear ranking function $f(\hat{x}) = \langle \theta, \hat{x} \rangle$ of this study is summarized based on all learned θ results from the 5-fold cross validation on “New”+“Rank”, where indices 1-8 in Fig. 8 denotes the left & right hippocampus, the left & right parahippocampal gyrus, the left & right putamen, and the left & right thalamus sequentially. It can be observed from Fig. 8 that, all of them have dominant influence on dementia disease diagnosis (i.e., all medians are within $[0.05, 0.15]$ without much difference), which complies well with clinical literatures in dementia studies [9, 11, 19].

5 Conclusion

In this study, a new dementia disease diagnosis strategy based on ASL-MRI is proposed for the first time. There are two steps composed of the whole strategy, including pixel-wise PVE correction and dementia disease severity prediction via ranking. Extensive experimental results and comprehensive statistical analysis demonstrate the superiority of the new disease diagnosis strategy. Main contributions of this study can be summarized as: 1) A new pixel-wise PVE correction method on ASL-MRI, which is capable to tackle problems of blurring and brain details loss in correction results commonly seen in conventional PVE correction methods as well as better improve CBF; 2) The first attempt to diagnose dementia disease based on corrected ASL-MRI from a new ranking perspective. Experimental analysis demonstrates the superiority of the newly proposed dementia disease diagnosis strategy over several compared existing methods. Future studies will be continued with more sophisticated ranking models investigated for severity prediction of diverse diseases reflected by different modalities of medical images.

Acknowledgments The authors would like to acknowledge national grants 61403182, 61363046, 61301194 and 61302121 approved by the National Natural Science Foundation of China, grants 20142BBE50023 and 20142BAB217033 approved by the Jiangxi Provincial Department of Science and Technology, as well as the NWPU grant 3102014JSJ0014 for supporting this study.

Appendix: Derivation of Equation (10)

After applying the chain rule, the gradient of C-NDCG(x) with respect to θ becomes:

$$\frac{\partial \text{C-NDCG}(x)}{\partial \theta} = \frac{\partial \text{C-NDCG}(x)}{\partial \pi'(x)} \cdot \frac{\partial \pi'(x)}{\partial \theta} = N_M^{-1} \sum_{x \in \chi} \frac{\partial \frac{2^{r(x)} - 1}{\log_2(1 + \pi'(x))}}{\partial \pi'(x)} \cdot \frac{\partial \pi'(x)}{\partial \theta} \tag{12}$$

where, the first term of (12) is derived as follows:

$$\frac{\partial \frac{2^{r(x)} - 1}{\log_2(1 + \pi'(x))}}{\partial \pi'(x)} = -\frac{2^{r(x)} - 1}{\log_2^2(1 + \pi'(x))} \cdot \frac{1}{(1 + \pi'(x)) \ln 2} \tag{13}$$

Furthermore, $\pi'(x)$ in (13) can be re-written as follows:

$$\pi'(x) \simeq 1 + \sum_{y \neq x, y \in \chi} \frac{s_{yx}}{\sqrt{s_{yx}^2 + \alpha^2}}, \quad s_{yx} = s_y - s_x \tag{14}$$

Apply the chain rule to the second term of (12) after incorporating results in (14):

$$\begin{aligned} \frac{\partial \pi'(x)}{\partial \theta} &= \frac{\partial \pi'(x)}{\partial s_{yx}} \cdot \frac{\partial s_{yx}}{\partial \theta} \\ &= \sum_{y \neq x, y \in \chi} \frac{\sqrt{s_{yx}^2 + \alpha^2} - s_{yx} \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{s_{yx}^2 + \alpha^2}} \cdot 2s_{yx}}{s_{yx}^2 + \alpha^2} \cdot \frac{\partial s_{yx}}{\partial \theta} \\ &= \sum_{y \neq x, y \in \chi} \frac{s_{yx}^2 + \alpha^2 - s_{yx} \cdot s_{yx}}{(s_{yx}^2 + \alpha^2)^{\frac{3}{2}}} \cdot \frac{\partial s_{yx}}{\partial \theta} \\ &= \sum_{y \neq x, y \in \chi} \frac{\alpha^2}{(s_{yx}^2 + \alpha^2)^{\frac{3}{2}}} \cdot \left(\frac{\partial f(\hat{y})}{\partial \theta} - \frac{\partial f(\hat{x})}{\partial \theta} \right) \end{aligned} \tag{15}$$

Hence, after substituting derivation results of (13) & (15) into (12), it becomes:

$$\begin{aligned} \frac{\partial \text{C-NDCG}(x)}{\partial \theta} &= N_M^{-1} \sum_{x \in \chi} \left(-\frac{2^{r(x)} - 1}{\log_2^2(1 + \pi'(x))} \cdot \frac{1}{(1 + \pi'(x)) \ln 2} \right) \\ &\quad \times \left(\sum_{y \neq x, y \in \chi} \frac{\alpha^2}{(s_{yx}^2 + \alpha^2)^{\frac{3}{2}}} \cdot \left(\frac{\partial f(\hat{y})}{\partial \theta} - \frac{\partial f(\hat{x})}{\partial \theta} \right) \right) \end{aligned} \tag{16}$$

which is the same as (10).

References

1. Asllani I, Borogovac A, Brown T (2008) Regression algorithm correcting for partial volume effects in arterial spin labeling MRI. *Magn Reson Med* 60:1362–1371
2. Brant-Zawadzki M, Gillan G, Nitz W (1992) MPRAGE: a three-dimensional, t1-weighted, gradient-echo sequence—initial experience in the brain. *Radiology* 182(3):769–775
3. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi M (2007) Forecasting the global burden of alzheimer's disease. *Alzheimers Dement* 3(3):186–191
4. Chen Y, Wolk D, Reddin J, Korczykowski M, Martinez P, Musiek E, Newberg A, Julin P, Arnold S, Greenberg J, Detre J (2011) Voxel-level comparison of arterial spin-labeling perfusion mri and fdg-pet in alzheimer disease. *Neurology* 77(2):1977–1985
5. Du Y, Tsui B, Frey E (2005) Partial volume effect compensation for quantitative brain spect imaging. *IEEE Trans Med Imaging* 24(8):969–976
6. Erlandsson K, Buvat I, Pretorius H, Thomas B, Hutton B (2012) A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol* 57(21):119–159
7. FMRIB Software Library (FSL) toolbox. <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>
8. Folstein M, Folstein S, McHung P (1975) Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12(3):189–198
9. Galton C, Patterson K, Graham K, Lambon-Ralph M, Williams G, Antoun N, Sahakian B, Hodges J (2001) Differing patterns of temporal atrophy in alzheimer's disease and semantic dementia. *Neurology* 57(2):216–225
10. Golay X, Hendrikse J, Lim T (2004) Perfusion imaging using arterial spin labeling. *Top Magn Reson Imaging: TMRI* 15(1):10–27
11. Gold G, Eniko K, Herrmann F, Canuto A, Hof P, Jean-Pierre M, Constantin B, Giannakopoulos P (2005) Cognitive consequences of thalamic, basal ganglia, and deep white matter lacunes in brain aging and dementia. *Stroke* 36(6):1184–1188
12. Goldstein T, Osher S (2008) The Split Bregman Method for L1 Regularized Problems. UCLA CAM report 08–29:1–21
13. Gunn R, Gunn S, Turkheimer F, Aston J, Cunningham V (2002) Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. *J Cereb Blood Flow Metab* 22:1425–1439
14. Individual Brain Atlases using Statistical Parametric Mapping (IBA-SPM) Software. <http://www.thomaskoenig.ch/Lester/ibaspm.htm>
15. Jarvelin K, Kekalainen J (2000) IR Evaluation Methods for Retrieving Highly Relevant Documents. In: *Proceedings ACM Special Interest Group Inf Retrieval (SIGIR)*, pp. 41–48
16. Joachims T SVM light - An Implementation of Support Vector Machine in C. URL <http://svmlight.joachims.org>
17. Johnson N, Jahng G, Weiner M, Miller B, Chui H, Jagust W, Gorno-Tempini M, Schuff N (2005) Pattern of cerebral hypoperfusion in alzheimer disease and mild cognitive impairment measured with arterial spin-labeling mr imaging: initial experience. *Radiology* 234:851–859
18. Keerthi S (2002) Efficient Tuning of SVM Hyperparameters using Radius/Margin Bound and Iterative Algorithms. *IEEE Trans Neural Netw* 13(5):1225–1229
19. Laakso M, Partanen K, Riekkinen P, Lehtovirta M, Helkala E, Hallikainen M, Hanninen T, Vainio P, Soininen H (1996) Hippocampal volumes in alzheimer's disease, parkinson's disease with and without dementia, and in vascular dementia an mri study. *Neurology* 46(3):678–681
20. Liu M, Zhang D, Shen D (2012) Ensemble sparse classification of alzheimers disease. *NeuroImage* 60(2):1106–1116
21. Mahendra B (1987) A pathography of dementia. *Dementia* 1:189–202
22. Malpass K, Disease Alzheimer (2012) Arterial Spin-Labeled MRI for Diagnosis and Monitoring of AD. *Nat Rev Neurol* 8(3):847–849

23. Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges J (2006) The addenbrooke's cognitive examination revised (ace-r): a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatr* 21(11):1078–1085
24. Murphy K, Brunberg J, Cohan R (1996) Adverse reactions to gadolinium contrast media: a review of 36 cases. *Am J Roentgenol* 167(4):847–849
25. Musiek E, Chen Y, Korczykowski M, Saboury B, Martinez P, Reddin J, Alavi A, Kimberg D, Wolk D, Julin P, Newberg A, Arnold S, Detre J (2012) Direct comparison of fluorodeoxyglucose positron emission tomography and arterial spin labeling magnetic resonance imaging in alzheimer's disease. *Alzheimers Dement* 8(1):51–59
26. Parkes L, Rashid W, Chard D, Tofts P (2004) Normal cerebral perfusion measurements using arterial spin labeling: reproducibility, stability, and age and gender effects. *Magn Reson Med* 51:736–743
27. Partial volume imaging. [http://en.wikipedia.org/wiki/Partial_volume_\(imaging\)](http://en.wikipedia.org/wiki/Partial_volume_(imaging))
28. Rice J (2007) *Mathematical Statistics and Data Dnalysis*, second edition. Duxbury Press
29. Statistical Parametric Mapping (SPM) toolbox. <http://www.fil.ion.ucl.ac.uk/spm/>
30. United Nations, World Population Prospects. <http://www.un.org/esa/population/publications/wpp2006/WPP2006.Highlights.rev.pdf>
31. Wang Z, Das S, Xie S, Arnold S, Detre J (2013) Arterial Spin Labeled MRI in Prodromal Alzheimer's Disease: A Multi-site Study. *NeuroImage: Clin* 2:630–636
32. Wee C, Yap P, Shen D (2013) Prediction of alzheimers disease and mild cognitive impairment using cortical morphological patterns. *Hum Brain Mapp* 34(12):3411–3425
33. World Health Organization, The Top 10 Causes of Death. <http://www.who.int/mediacentre/factsheets/fs310/en/index2.html>
34. Wright M (2004) The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bull Am Math Soc* 42(39):1–16
35. Zhou L, Wang Y, Li Y, Yap P, Shen D (2011) Hierarchical Anatomical Brain Networks for MCI Prediction: Revisiting Volumetric Measures. *PLoS Comput Biol* 6(7):e21935



Wei Huang obtained his B.Eng. and M.Eng. from the Harbin Institute of Technology, China in 2004 and 2006, respectively. He obtained his Ph.D. from the Nanyang Technological University, Singapore in 2011. Before joining the Nanchang University as an Associate Professor, he worked in the University of California San Diego, USA, as well as the Agency for Science Technology and Research, Singapore as Research Staffs. Dr. Huang's recent research interests mainly include but not limited to medical image processing, pattern recognition, and computer vision.



Peng Zhang received the B.E. degree from the Xi'an Jiaotong University, China in 2001. He received his Ph.D. from Nanyang Technological University, Singapore in 2011. He is now an Associate Professor in School of Computer Science, Northwestern Polytechnical University, China. His current research interests include image processing, signal processing, and computer vision. He is a member of IEEE. He is the TPC member and reviewer of several high ranked international conferences and journals.



Minmin Shen received the B.Sci. degree in information engineering from the Shanghai Jiao Tong University, Shanghai, China, in 2006, and the Ph.D. degree from the Nanyang Technological University, Singapore, in 2011. She was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University from 2010 to 2011. She is currently a Marie Curie Fellow at the Department of Computer and Information Science, University of Konstanz, Germany. She is also with School of Software Engineering, South China University of Technology. Her current research interests include biological image/video processing, computer vision, super-resolution, video coding, and video/image enhancement.