DEEP LEARNING & NEURAL COMPUTING FOR INTELLIGENT SENSING AND CONTROL

Classification of Alzheimer's disease based on brain MRI and machine learning

Zhao Fan[1] · Fanyu Xu[2] · Xuedan Qi[3] · Cai Li[4] · Lili Yao[4]

Abstract

Alzheimer's disease (AD) is one of the most common diseases in the world. It is a neurodegenerative disease that can cause cognitive impairment and memory deterioration. In recent years, the number of the elderly population is increasing, and the incidence of elderly diseases has increased significantly. The most representative of these diseases is Alzheimer's disease. According to some data, the average survival time of Alzheimer's disease patients is only 5.5 years, which is the "fourth killer" that endangers the health of the elderly after cardiovascular diseases, cerebrovascular diseases and cancer. According to conservative estimates of the International Federation of Alzheimer's Diseases, the number of Alzheimer's disease patients worldwide will increase to 75.62 million by 2030; by 2050, the number of patients will reach 135.46 million. Therefore, it is urgent to classify the course of Alzheimer's disease. In this paper, support vector machine (SVM) model method is used to classify and predict different disease processes of Alzheimer's disease based on structural brain magnetic resonance imaging (MRI) imaging data, so as to help the auxiliary diagnosis of the disease. In this paper, the extracted MRI data and the SVM model are combined to obtain more accurate classification prediction results. The accuracy of classification and prediction is the best. According to the predicted results, the data characteristics related to diseases can be determined, which can provide a basis for clinical and basic research, etiology and pathological changes.

Keywords Machine learning · Support vector machine · Brain MRI · Alzheimer's disease

1 Introduction

Alzheimer's disease (AD) is the most common and a neurodegenerative disease with hidden onset and progressive development [1–3] accounting for 50–60% of all patients. Alzheimer's disease can lead to changes in memory, cognitive function and behavior. The damage to the brain is usually irreversible. After about 3–10 years, the patient will eventually die [4]. This disease has gradually attracted people's attention and become a hot issue in society. According to the development of cognitive model and the degree of impaired function, the onset of Alzheimer's disease can be divided into four stages: normal (control), mild (incipient), moderate (moderate) and severe

✉ Zhao Fan
fanzhao316@sxmu.edu.cn

Fanyu Xu
201701030962@b.sxmu.edu.cn

Xuedan Qi
041510520@b.sxmu.edu.cn

Cai Li
201700410564@b.sxmu.edu.cn

Lili Yao
201800410514@b.sxmu.edu.cn

[1] Institute of Geriatrics, Shanxi Medical University, Taiyuan 030001, China

[2] School of Stomatology, Shanxi Medical University, Taiyuan 030001, China

[3] 306 Hospital of PLA, Beijing 100101, China

[4] Department of Physiology, Basic Medical Sciences of Shanxi Medical University, Taiyuan 030001, China

(severe) [5]. Normal stage is the early stage of Alzheimer's disease, patients will have mild cognitive difficulties and memory loss, the most common symptom is the lack of ability to access new information, it is difficult to recall recent events, and many complex daily activities will also be affected [6]. The learning and memory impairment of patients in mild stage are further increased. Some patients have more difficulties in language, executive function and perception than in memory. Alzheimer's disease does not affect all memory content, such as situational memory, implicit memory and semantic memory, but has a lower impact on the memory of new things. Alzheimer's disease patients at this stage can generally express their basic views, but they may have difficulties in writing, painting, decoration, planning and some coordinated actions. In the moderate stage, the patient's condition deteriorated further, even affecting the patient's independence, eventually making them unable to carry out basic daily activities. At this stage, patients do not have the ability to memorize vocabulary, language barriers become obvious, language disorders often occur, and reading and writing skills gradually disappear [7–9]. Patients may have symptoms of not knowing their relatives and friends around them, and complete long-term memory begins to damage. The common clinical manifestations are emotional instability, irritability and violence. Severe stage is the late stage of Alzheimer's disease, patients rely entirely on caregivers. Expressive ability degenerates to simple phrases and sub-words, and ultimately language function is completely lost [10]. With the aging of the global population and the increasing number of dementia patients, AD has become a major social problem we are facing. According to the WHO report in 2009, there are 17.8 million AD patients in the world, which is expected to reach 57.5 million by 2050. In the USA, AD has posed a threat to human life and health, becoming the fourth leading death killer after cardiovascular disease, cancer and brain death. In the twenty-first century, the world will be an irreversible aging society. The absolute number of the elderly is quite large. By the year 2005, the world's aging population will reach 1.837 billion. It can be predicted that with the development of global aging society, AD will seriously affect the development of countries all over the world [11, 12]. Therefore, it is very meaningful to study AD disease.

At present, a large number of researchers at home and abroad have studied the functional imaging and structural imaging of the brain of AD patients and then analyzed what changes will occur in the brain structure, blood flow and metabolism when memory, thinking and behavior disorders occur in AD patients, so as to find a biomarker for clinical diagnosis [13–15]. Recently, a voxel-based neuroimaging study found that the volume of gray matter in bilateral medial temporal lobe, left medial thalamus, insula lobe and left middle temporal gyrus/superior temporal gyrus decreased significantly in AD patients compared with the control group [16]. In addition, Wang et al. [17] found that only the left medial temporal lobe had more severe gray matter atrophy in AD patients than MCI patients. They suggested that the deterioration of left medial temporal lobe atrophy could predict the transformation of MCI to AD [18]. To some extent, the left medial temporal lobe could be used as a monitor for the progression of the disease. In the study of functional imaging, GPrestia A concluded that when AD patients performed situational memory, the results of functional imaging showed that during encoding and extraction, AD patients showed negative activation in the medial temporal lobe and positive activation in the ventral prefrontal cortex. Scholars Peskind et al. pointed out in the study that [19] specific activation areas appeared in the left middle frontal gyrus dorsal and left prefrontal lobe of AD patients. However, the current studies at home and abroad are all single-mode analysis, focusing only on brain functional imaging or structural imaging. The results of the study are different because of the different experimental conditions. No researchers have used this method to carry out multi-mode meta-analysis of brain structural and functional changes in AD patients.

In addition to providing clinical diagnostic value by studying the changes of structure and function of brain neuroimaging in AD patients [20], structural MR images have the advantages of high resolution, simple operation and low cost for soft tissue and can clearly distinguish gray matter from white matter without artifact interference of skull base and can display brain structure more clearly. Meanwhile, texture analysis technology is also used. Studying the change rule and distribution pattern of gray value of image pixels can sensitively reflect the subtle changes of tissue pathology. Amlerova et al. used wavelet analysis to identify temporal lobe epilepsy with a correct rate of 94% [21], while the correct rate of morphological recognition was 83%. Liu et al. recognized AD, MCI and NC through three-dimensional texture analysis technology [22, 23]. The results also showed that there were significant differences in texture parameters between patients and control groups, and the recognition results by texture analysis reached 83.33%. These studies tell us that the method of classifying diseases through the intrinsic information of images may be more helpful for clinical diagnosis. Principal component analysis (PCA) is one of the most widely used feature extraction methods at present. It can extract the main information in images. It has been applied in various signal processing-related research fields [24, 25]. Liu et al. used PCA to realize face recognition system [26–28]. Therefore, this paper also attempts to use principal component analysis combined with linear

discriminant analysis (LDA) and support vector machine (SVM) to classify and recognize AD, MCI and NC.

In this paper, support vector machine (SVM) model method is used to classify and predict different disease processes of Alzheimer's disease based on structural brain magnetic resonance imaging (MRI) imaging data, so as to help the auxiliary diagnosis of the disease. In this paper, the extracted MRI data and the SVM model are combined to obtain more accurate classification prediction results. The accuracy of classification and prediction is the best. According to the predicted results, the data characteristics related to diseases can be determined, which can provide a basis for clinical and basic research, etiology and pathological changes.

2 Proposed method

2.1 Machine learning

Machine learning (ML) is a multi-disciplinary interdisciplinary subject, involving probability theory, statistics, approximation theory, convex analysis, algorithm complexity theory and other disciplines. It specializes in how computers simulate or implement human learning behavior to acquire new knowledge or skills, reorganize existing knowledge structures and continuously improve their own performance. It is the core of artificial intelligence and the fundamental way to make computers intelligent. It is applied in all fields of artificial intelligence. It mainly uses induction and synthesis rather than deduction.

2.1.1 Main strategies and basic structure of machine learning

1. Main strategies

Learning is a complex intelligent activity. Learning process is closely related to reasoning process. According to the amount of reasoning used in learning, machine learning strategies can be generally divided into four categories: mechanical learning, learning through imparting, analogy learning and learning through examples. The more the reasoning used in learning, the stronger the ability of the system.

2. Basic structure

It represents the basic structure of the learning system. The environment provides some information to the learning part of the system. The learning part uses this information to modify the knowledge base in order to improve the efficiency of the execution part of the system in accomplishing tasks. The execution part completes tasks according to the knowledge base and feeds back the information obtained to the learning part. In the specific application, the environment, knowledge base and execution part determine the specific work content, and the problems to be solved in the learning part are completely determined by the above three parts. Next, we describe the impact of these three parts on the design of learning systems.

The most important factor affecting the design of learning system is the information provided by the environment to the system or, more specifically, the quality of information. The knowledge base contains general principles that guide the action of the executive part, but the information provided by the environment to the learning system is different. If the quality of information is relatively high and the difference with general principles is relatively small, then the learning part is easier to deal with. If the specific information provided to the learning system is cluttered disorderly to guide the implementation of specific actions, the learning system needs to delete unnecessary details, summarize and popularize them, form general principles for guiding actions, and put them into the knowledge base after obtaining sufficient data. Thus, the task of the learning part is more onerous and difficult to design.

Because the information acquired by the learning system is often incomplete, the reasoning conducted by the learning system is not entirely reliable, and the rules it summarizes may be correct or incorrect. This should be tested by the effect of implementation. Correct rules can improve the efficiency of the system and should be retained; incorrect rules should be modified or deleted from the database.

Knowledge base is the second factor affecting the design of learning system. There are many forms of knowledge representation, such as eigenvectors, first-order logic statements, production rules, semantic networks and frameworks. These representations have their own characteristics. When choosing the representations, the following four aspects should be taken into account:

1. Strong expression ability.
2. Easy to reason.
3. It is easy to modify the knowledge base.
4. Knowledge representation is easy to expand.

One of the last issues to be addressed in the knowledge base is that the learning system cannot acquire knowledge without any knowledge at all. Each learning system requires information provided by certain knowledge understanding environments, analysis and comparison, making assumptions, testing and modifying these assumptions. Therefore, more precisely, the learning

system is an extension and improvement of existing knowledge.

2.1.2 Classification

1. Classification based on learning strategies

Learning strategy refers to the reasoning strategy adopted by the system in the process of learning. The classification criteria of learning strategies are classified according to the degree of reasoning and difficulty that students need to achieve information conversion. They are divided into six basic types in the order of simplicity to complexity and from fewer to more:

(1) Rote learning

Without any reasoning or other knowledge conversion, learners can directly absorb the information provided by the environment, such as Samuel's checkers program, Newell and Simon's LT system. This kind of learning system mainly considers how to index and utilize the stored knowledge. The systematic learning method is to learn directly through pre-programmed and constructed procedures. The learners do not do any work, or learn directly by receiving the established facts and data, and do not make any reasoning about the input information.

(2) Learning by deduction

Reasoning proceeds from axioms and deduces conclusions through logical transformation. This reasoning is a process of "fidelity" transformation and specialization, which enables students to acquire useful knowledge in the reasoning process. This learning method includes macro-operation learning, knowledge editing and Chunking technology. The inverse process of deductive reasoning is inductive reasoning.

(3) Learning by analogy

By using the similarity of knowledge in two different domains (source domain and target domain), we can deduce the corresponding knowledge of target domain from the knowledge of source domain (including similar characteristics and other properties) by analogy, so as to realize learning. Analogy learning system can make an existing computer application system adapt to new fields to complete similar functions that were not designed before.

Analogical learning requires more reasoning than the three learning methods mentioned above. It generally requires that the available knowledge be retrieved from the knowledge source (source domain) and then transformed into a new form for use in the new situation (target domain). Analogy learning plays an important role in the history of human science and technology. Many scientific

discoveries are obtained by analogy. For example, the famous Rutherford analogy reveals the mystery of atomic structure by analogizing the atomic structure (target domain) with the solar system (source domain).

(4) Learning from induction

The reasoning workload of this kind of learning is higher than that of demonstration learning and deductive learning, because the environment does not provide general conceptual descriptions (such as axioms). To some extent, inductive learning has a larger amount of reasoning than analogical learning, because no similar concept can be used as a "source concept." Inductive learning is the most basic and mature learning method, which has been widely studied and applied in the field of artificial intelligence.

2. Comprehensive classification

Considering the historical origin, knowledge representation, reasoning strategy, similarity of result evaluation, relative concentration of researchers' communication and application fields of various learning methods, machine learning methods are divided into the following six categories:

(1) Empirical inductive learning

Empirical inductive learning uses some data-intensive empirical methods (such as version space method, ID3 method, law discovery method) to conduct inductive learning of examples. Its examples and learning results are generally represented by attributes, predicates, relations and other symbols. It is equivalent to inductive learning based on the classification of learning strategies, but deducts the parts of link learning, genetic algorithm and reinforcement learning.

(2) Analytic learning

Analytical learning is based on one or a few examples, using domain knowledge for analysis. Its main features are: reasoning strategy is mainly deduction, not induction; using past problem solving experience (examples) to guide new problem solving, or generating search control rules that can more effectively use domain knowledge. The goal of analytical learning is to improve the performance of the system, not to describe new concepts. Analytical learning includes application interpretation learning, deductive learning, multi-level structure block learning and macro-operation learning.

(3) Analogical Learning

It is equivalent to analogical learning based on classification of learning strategies. The most striking research in this type of learning is the concrete example of learning

past experience through analogy, which is called case-based learning or abbreviation-based case-based learning.

(4) Genetic algorithm

Genetic algorithm simulates mutation, exchange of biological reproduction and Darwin's natural selection (survival of the fittest in every ecological environment). It codes the possible solution of the problem as a vector, called an individual, and each element of the vector is called a gene. It evaluates each individual in a population (a set of individuals) by using the objective function (corresponding to the natural selection criteria) and carries out genetic operations such as selection, exchange and mutation according to the evaluation value (fitness), so as to obtain a new population. Genetic algorithm is suitable for very complex and difficult environments, such as with a lot of noise and irrelevant data, things are constantly updated, problem objectives cannot be clearly and accurately defined, and the value of current behavior can be determined through a long execution process. Like neural networks, genetic algorithm has developed into an independent branch of artificial intelligence, whose representative is Holland [29].

(5) Linked Learning

A typical connection model is an artificial neural network, which consists of some simple computational units called neurons and weighted connections between units.

(6) Reinforcement learning

Enhanced learning is characterized by determining and optimizing action choices through trial-and-error interaction with the environment to achieve so-called sequential decision-making tasks. In this task, the learning mechanism can change the state of the system by choosing and executing actions and may get some reinforcement signal (immediate return), thus realizing the interaction with the environment. Enhanced signal is a kind of scalar reward and punishment for system behavior. The goal of system learning is to find an appropriate action selection strategy, that is, the method of choosing which action in any given state, so that the generated action sequence can obtain some optimal results (e.g., the maximum cumulative immediate return).

In comprehensive classification, empirical inductive learning, genetic algorithm, join learning and reinforcement learning belong to inductive learning. Among them, empirical inductive learning adopts symbolic representation, while genetic algorithm, join learning and reinforcement learning adopt sub-symbolic representation; analytical learning belongs to deductive learning. In fact, the strategy of analogy can be regarded as the synthesis of inductive and deductive strategies. Therefore, the most

basic learning strategies are induction and deduction. From the point of view of learning content, inductive learning strategy can be called knowledge-level learning because it induces input, and the knowledge learned is obviously beyond the scope of the original system knowledge base. The results of learning change the knowledge deduction closure of the system, so this type of learning can be called knowledge-level learning. While the knowledge learned by using deductive strategy can improve the efficiency of the system, it can also be called knowledge-level learning. It can still be contained in the knowledge base of the original system, that is, the knowledge learned has not changed the deductive closure of the system, so this type of learning is also called symbolic learning.

2.1.3 Classification of learning forms

1. Supervised learning

Supervised learning is to provide right and wrong instructions in the process of mechanical learning. Generally speaking, the final result (0, 1) is contained in the data group. Through the algorithm, the machine can reduce the error by itself. This kind of learning is mainly applied to classification and prediction (regression and classify). Supervised learning learns a function from a given training data set, which can be used to predict the results when new data arrive. The training set of supervised learning requires input and output, or features and objectives. The goal of training concentration is marked by people. Common supervised learning algorithms include regression analysis and statistical classification.

2. Unsupervised learning

Unsupervised learning, also known as inductive learning (clustering), uses K-means to establish a centered and reduces errors through iteration and descent operations to achieve the purpose of classification.

2.2 Support vector machine

Vapnik et al. put forward another optimal design criterion for linear classifiers based on statistical learning theory for many years. The principle also starts with linear separability and then extends to the case of linear inseparability. Even extended to the use of nonlinear functions, this classifier is called support vector machine (SVM). Support vector machine (SVM) has a deep theoretical background. Support vector machine (SVM) is a new method proposed in recent years.

2.2.1 Basic ideas of support vector machine

The main idea of SVM can be summarized as two points:

1. It analyzes the linear separable case. For the linear inseparable case, the linear inseparable sample in the low-dimensional input space is transformed into the high-dimensional feature space by using the nonlinear mapping algorithm, which makes it possible for the high-dimensional feature space to use the linear algorithm to analyze the nonlinear feature of the sample.

2. Based on the theory of structural risk minimization, it constructs an optimal partitioning hyperplane in feature space, so that the learner can be globally optimized, and the expected risk in the whole sample space satisfies a certain upper bound with a certain probability. In learning this method, we should first understand the characteristics of the problem considered in this method, which should be discussed from the simplest case of linear separability. Before we understand its principle, we should not rush to learn the more complicated cases of linear inseparability. In designing support vector machines, we need to use the solution of conditional extremum problem, so we need to use Lagrange multiplier theory, but for most cases. For people, what they have learned before or used commonly is that the constraints are expressed in terms of equality, but inequalities are used here as conditions that must be satisfied. At this time, it is only necessary to understand the relevant conclusions of Lagrange's theory.

2.2.2 Basic principles of SVM

The SVM method maps the sample space into a high-dimensional or even infinite-dimensional feature space (Hilbert space) through a nonlinear mapping p, which transforms the nonlinear separable problem in the original sample space into a linear separable problem in the feature space. Dimension-raising means mapping samples to high-dimensional space. Generally, it will increase the complexity of computation and even cause "dimension disaster," so people seldom ask for it. However, as a classification and regression problem, it is likely that the sample set which cannot be linearly processed in low-dimensional sample space can be linearly partitioned (or regressed) by a linear hyperplane in high-dimensional feature space. Generally, dimensionality updating will lead to computational complexity. SVM ingeniously solves this problem: Applying the expansion theorem of kernel function, we do not need to know the explicit expression of nonlinear mapping; because linear learning machine is

built in high-dimensional feature space, compared with linear model, it not only hardly increases computational complexity, but also avoids "dimension" to some extent. All this is attributed to the expansion and computational theory of kernels.

The following describes the two-category problem: Consider a set of l sample points in the n_1-dimensional space:

$$T = \{(x_1, y_1), \ldots, (x_l, y_l)\} \tag{1}$$

$x_i \in R^n$ is the input vector and $y_i \in \{-1, 1\}$ is the class label of x_i. According to the given training set T, a real-valued function $g(x)$ on R^n is searched. The decision function $f(x) = \text{sgn}(g(x))$ is used to deduce the corresponding y for any given x. If there is a classification hyperplane,

$$w^T \cdot x + b = 0. \tag{2}$$

If the same class of samples can accurately fall on the same side of the classification hyperplane, then the sample set is called linear separable. Namely satisfy

$$\begin{cases} w \cdot x_i + b \geq 1, y_i = 1 \\ w \cdot x_i + b \leq -1, y_i = -1 \end{cases} \quad i = 1, 2, \ldots, l \tag{3}$$

Define the interval between sample point x_i and classification hyperplane as follows:

$$\varepsilon_i = y_i(w \cdot x_i + b) = |w \cdot x_i + b| \tag{4}$$

The w and b in the above formula are normalized, replaced by $\frac{w}{\|w\|}$ and $\frac{b}{\|w\|}$, and the interval is defined as geometric interval.

$$\delta_i = \frac{w \cdot x_i + b}{\|w\|} \tag{5}$$

On this basis, the distance from a sample set to the classification hyperplane is defined as the geometric distance between the nearest sample point in the set and the classification hyperplane, that is, the distance between the sample set and the classification hyperplane.

$$\delta = \min \delta_i, \quad i = 1, 2, \ldots, l \tag{6}$$

Therefore, an optimal classification plane is selected among the innumerable classification hyperplanes satisfying the formula (3), so that the distance δ of the sample set to the classification hyperplane is the largest. When $\varepsilon_i = |w \cdot x_i + b| = 1$, the distance between the two samples is

$$2\frac{|w \cdot x_i + b|}{\|w\|} = \frac{2}{\|w\|}. \tag{7}$$

The goal is to find the optimal classification hyperplane to maximize $\frac{2}{\|w\|}$ and minimize $\frac{\|w\|^2}{2}$ under the constraint of

formula 3. Convert this optimization problem into a mathematical language as follows, that is

$$\begin{cases} \min \frac{\|w\|^2}{2} \\ s.t. y_i(wx_i+b) \ge 1, \end{cases} \quad i = 1, 2, \ldots, l \qquad (8)$$

The convex quadratic programming problem with variables w and b is called the primitive problem. For this reason, the primitive problem is not solved directly, but its solution is obtained through its dual problem.

3 Experiments

3.1 Data sources

The data in this study are from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. ADNI was launched in 2003 by the National Institute on Ageing (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations. It invested \$60 million to recruit large numbers of subjects aged 55 to 90 from more than 50 sites in the USA and Canada, including elderly people with normal cognitive abilities: human, early and late MCI population, AD patients, through the examination of MRI, PET and other biological markers, combined with clinical and neuropsychological assessment, to detect the MCI process and the large database of early AD.

3.2 Experimental steps

1. In Matlab version 7.11.0 software, empty environment variables and import data. According to LIBSVM requirements, the data are sorted out, and rows represent samples, and each row represents a research object. Each column is a feature.
2. Constructing classification model. The data samples are randomly divided into training set samples and test set samples. The SVM classifier is trained with the training set sample data. Thus, a classification model is obtained, and the classification of test set samples is evaluated according to the rules of training and learning. In this study, the svmtrain function of the software package is used to construct the classifier, obtain the classification model, constantly change the training set and test set sample data until all data are predicted and cross-validated to find the optimal parameters.
3. Classification test. The test set samples are classified in the trained model to obtain the predictive classification results and the classification accuracy. The average of the classification prediction accuracy of all the data is

taken to represent the predictive accuracy of the classifier, and the simulation test chart of SVM is given. Svmpredict function was used for classification prediction in this study.

4 Discussion

4.1 Comparison of classification results of different measurement features

Different measurement features (subcortical volume, cortical volume, cortical thickness, surface area and different volume of hippocampal structure) were put into the model one by one, that is, each feature was trained and tested as the only feature. The classification accuracy of cortical surface area (SA) as the only parameter was 51.44%. The classification accuracy of HS as the only parameter was 63.78%. The classification accuracy of TA was 56.37%. The classification accuracy of SV was 55.96%. The classification accuracy of CV was 56.79%. Comparison of classification results for different measurement features is shown in Fig. 1.

4.2 Comparison of classification results of samples from different training sets

Among 317 sample training sets, 230 test sets had different prediction accuracy rates; 281 test sets had 48.61% classification accuracy; 56 test sets had 97.87% classification accuracy; 9 test sets had 65.49% classification accuracy; 142 test sets had 59.30% classification accuracy; 47 test sets had 76.73% classification accuracy. The sample classification accuracy of item index test set is 65.49%, as shown in Fig. 2.

Among 688 sample training sets, 151 test sets had different prediction accuracy rates; 281 test sets had a classification accuracy of 55.42%; 56 test sets had a
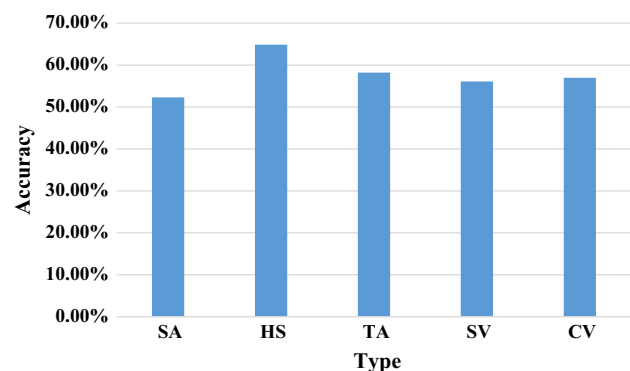

Fig. 1 Comparison of classification results for different measurement features
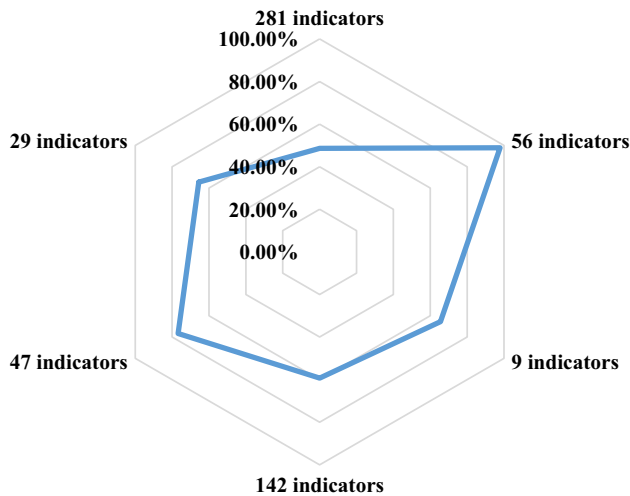
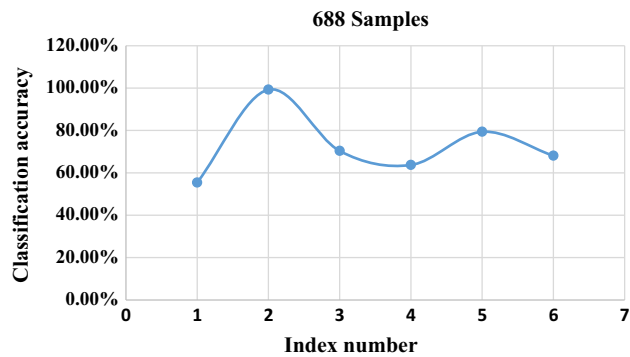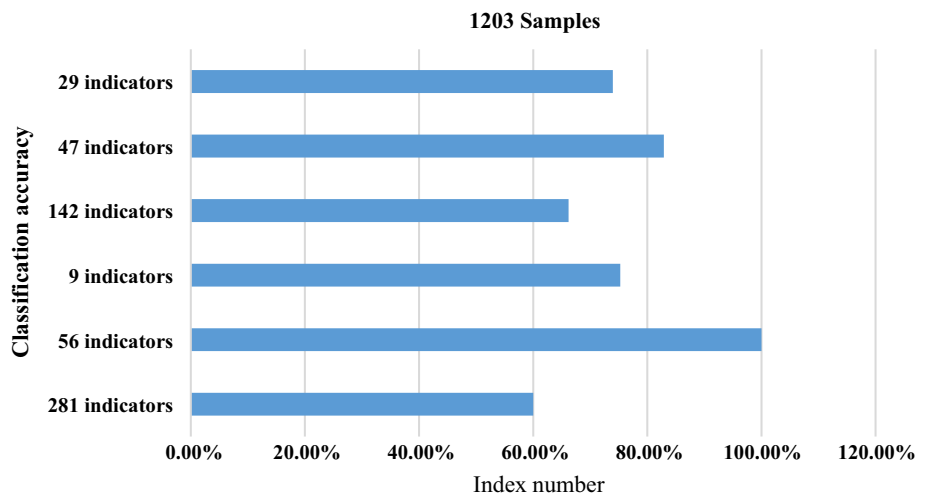Fig. 2 Classification results of 311 samples with different indicators

Fig. 3 Classification results of 688 samples with different indicators

classification accuracy of 99.27%; 9 test sets had a classification accuracy of 70.35%; 142 test sets had a classification accuracy of 63.71%; 47 test sets had a classification accuracy of 79.38%. The sample classification accuracy of item index test set is 68.07%, as shown in Fig. 3.

Among 1203 sample training sets, 69 test sets had different prediction accuracy; 281 test sets had 60.02% classification accuracy; 56 test sets had 100% classification accuracy; 9 test sets had 75.26% classification accuracy; 142 test sets had 66.21% classification accuracy; 47 test sets had 82.90% classification accuracy. The sample classification accuracy of item index test set is 73.97%, as shown in Fig. 4.

Combining the above results, the prediction results of different indicators in different training set samples are compared as shown in Fig. 5.

5 Conclusions

1. This paper introduces the basic concepts, classification and learning forms of machine learning. At the same time, the basic idea and principle of support vector machine are introduced.

2. In this paper, the most popular classification method, support vector machine (SVM), is used to predict the course of AD disease. Based on the multi-dimensionality of decision space, it does not need a large number of redundant samples and complex data preprocessing, and the algorithm is simple and has good robustness. The model established by SVM method has strong objectivity and good generalization ability. This study has shown the superiority of SVM classifier in predicting different stages of AD. It shows that SVM method has a strong ability to fit the disease process of AD and can be used to evaluate unknown samples. We are collecting relevant clinical cases and combining with the actual situation of the Chinese population, carrying out further analysis and application. The parameters of image data scanning in different hospitals are processed according to the need. On this basis,

Fig. 4 Classification results of 1203 samples with different indicators

Fig. 5 Comparison of classification results of different indicators in different training sets

Comparisons of the classification accuracy of the number of samples in different test sets
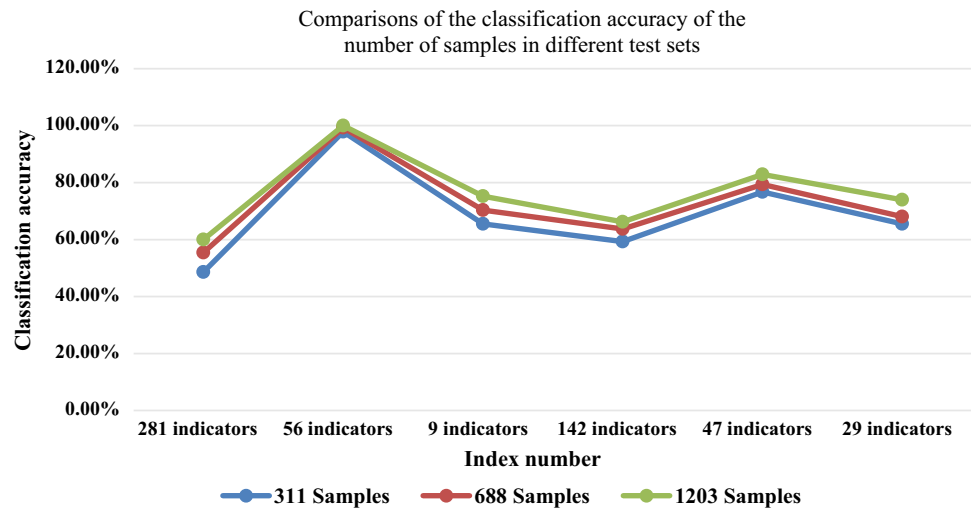
a model is constructed for training and prediction to accurately predict the conversion of MCI to AD, and corresponding prevention and treatment strategies are formulated to provide a warning line for patients' families and clinicians.

3. In this paper, we extract the characteristics of brain MRI of different groups of people and construct SVM model to predict the different processes of AD, and achieve better results. The data source of ADNI database used in this paper is accurate and reliable, and the method of index extraction and compression is appropriate. In addition, the data volume of training samples may be too small, so that the accuracy of classification and prediction of training samples can reach 100%.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Reitz C, Mayeux R (2014) Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. Biochem Pharmacol 88(4):640–651
2. Green RC, Cupples LA, Kurz A et al (2016) Depression as a risk factor for Alzheimer disease: the MIRAGE study. Arch Neurol 60(5):753
3. Bloom GS (2014) Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol 71(4):505–508
4. Lucía CG, Leen B, Iryna B et al (2014) The mechanism of γ-Secretase dysfunction in familial Alzheimer disease. EMBO J 31(10):2261–2274
5. Tarasoffconway JM, Carare RO, Osorio RS et al (2016) Clearance systems in the brain—implications for Alzheimer disease. Nat Rev Neurol 12(4):248
6. Ferrucci R, Mameli F, Guidi I et al (2016) Transcranial direct current stimulation improves recognition memory in Alzheimer disease. Neurology 71(7):493–498
7. Ujiie M, Dickstein DL, Carlow DA et al (2015) Blood-brain barrier permeability precedes senile plaque formation in an Alzheimer disease model. Neurobiol Aging 25(6):S236–S236
8. Barker WW, Luis CA, Kashuba A et al (2015) Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the State of Florida Brain Bank. Alzheimer Dis Assoc Disord 16(4):203–212
9. Tariot PN, Farlow MR, Grossberg GT et al (2014) Memantine treatment in patients with moderate to severe Alzheimer disease already receiving donepezil: a randomized controlled trial. Chin J Gen Pract 291(3):317–324
10. Tromp D, Dufour A, Lithfous S et al (2015) Episodic memory in normal aging and Alzheimer disease: insights from imaging and behavioral studies. Ageing Res Rev 24(Pt B):232–262
11. Butterfield DA, Domenico FD, Barone E (2014) Elevated risk of type 2 diabetes for development of Alzheimer disease: a key role for oxidative stress in brain. Biochim Biophys Acta 1842(9):1693–1706
12. Rd DD III, Jerskey BA, Chen K et al (2014) Brain differences in infants at differential genetic risk for late-onset Alzheimer disease: a cross-sectional imaging study. JAMA Neurol 71(1):11
13. Takeda S, Sato N, Morishita R (2014) Systemic inflammation, blood-brain barrier vulnerability and cognitive/non-cognitive symptoms in Alzheimer disease: relevance to pathogenesis and therapy. Front Aging Neurosci 6(6):171
14. Willette AA, Bendlin BB, Starks EJ et al (2015) Association of insulin resistance with cerebral glucose uptake in late middle-aged adults at risk for Alzheimer disease. JAMA Neurol 72(9):1013
15. Kester MI, Goos JD, Teunissen CE et al (2014) Associations between cerebral small-vessel disease and Alzheimer disease pathology as measured by cerebrospinal fluid biomarkers. JAMA Neurol 71(7):855–862
16. Tramutola A, Triplett JC, Domenico FD et al (2015) Alteration of mTOR signaling occurs early in the progression of Alzheimer disease (AD): analysis of brain from subjects with pre-clinical

AD, amnestic mild cognitive impairment and late-stage AD. J Neurochem 133(5):739–749

17. Wang LY, Raskind MA, Wilkinson CW et al (2018) Associations between CSF cortisol and CSF norepinephrine in cognitively normal controls and patients with amnestic MCI and AD dementia. Int J Geriatr Psychiatry 33(5):763

18. Prestia A, Caroli A, Wade SK et al (2015) Prediction of AD dementia by biomarkers following the NIA-AA and IWG diagnostic criteria in MCI patients from three European memory clinics. Alzheimers Dement J Alzheimers Assoc 11(10):1191–1201

19. Peskind ER, Tsuang DW, Bonner LT et al (2015) Propranolol for disruptive behaviors in nursing home residents with probable or possible Alzheimer disease: a placebo-controlled study. Alzheimer Dis Assoc Disord 19(1):23–28

20. Carnevale L, D'Angelosante V, Landolfi A et al (2018) Brain MRI fiber-tracking reveals white matter alterations in hypertensive patients without damage at conventional neuroimaging. Cardiovas Res 114(11):1536–1546

21. Amlerova J, Cavanna AE, Bradac O et al (2014) Emotion recognition and social cognition in temporal lobe epilepsy and the effect of epilepsy surgery. Epilepsy Behav 36(36):86–89

22. Liu WF, Wang X, Xia H (2014) 3D texture analysis of Corpus Caliosum based on MR images inpatients with Alzheimer's disease and mild cognitive impairment. Appl Mech Mater 533:415–420

23. Yu L, Xia H, Liu W (2016) Classification studies in patients with Alzheimer's disease and normal control group based on three-dimensional texture features of hippocampus magnetic resonance images. J Biomed Eng 33(6):1090–1094

24. Chen Z, Chen X, Chen Z et al (2017) Alteration of gray matter texture features over the whole brain in medication-overuse headache using a 3-dimensional texture analysis. J Headache Pain 18(1):112

25. Leandrou S, Petroudi S, Kyriacou PA et al (2018) Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: a methodological review. IEEE Rev Biomed Eng 11(99):97

26. Liu Z, Lv L, Yong W (2017) Development of face recognition system based on PCA and LBP for intelligent anti-theft doors. In: IEEE International conference on computer and communications

27. Wang Q (2016) Design and implementation of remote facial expression recognition surveillance system based on PCA and KNN algorithms. In: International conference on intelligent information hiding and multimedia signal processing

28. Guo-Jun MA, Zhou HD (2015) Research and implementation of intelligent terminal lightweight face recognition system. J Commun 5(6):119–134

29. Liang X (2019) A medical diagnostic expert system with LPSO-based artificial neural network. Investigación Clínica 60(4):894–899

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.