

## Research Article

# A Hybrid Intelligent Diagnosis Approach for Quick Screening of Alzheimer's Disease Based on Multiple Neuropsychological Rating Scales

Ziming Yin,<sup>1</sup> Yinhong Zhao,<sup>2</sup> Xudong Lu,<sup>1</sup> and Huilong Duan<sup>1</sup>

<sup>1</sup>College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, Zhejiang 31002, China

<sup>2</sup>China National Center for Biotechnology Development, Building D, No. 16, Xisihuanzhonglu, Haidian District, Beijing 100036, China

Correspondence should be addressed to Xudong Lu; [lxldxd66@gmail.com](mailto:lxldxd66@gmail.com)

Received 6 June 2014; Revised 20 November 2014; Accepted 20 November 2014

Academic Editor: José M. Jerez

Copyright © 2015 Ziming Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neuropsychological testing is an effective means for the screening of Alzheimer's disease. Multiple neuropsychological rating scales should be used together to get subjects' comprehensive cognitive state due to the limitation of a single scale, but it is difficult to operate in primary clinical settings because of the inadequacy of time and qualified clinicians. Aiming at identifying AD's stages more accurately and conveniently in screening, we proposed a computer-aided diagnosis approach based on critical items extracted from multiple neuropsychological scales. The proposed hybrid intelligent approach combines the strengths of rough sets, genetic algorithm, and Bayesian network. There are two stages: one is attributes reduction technique based on rough sets and genetic algorithm, which can find out the most discriminative items for AD diagnosis in scales; the other is uncertain reasoning technique based on Bayesian network, which can forecast the probability of suffering from AD. The experimental data set consists of 500 cases collected by a top hospital in China and each case is determined by the expert panel. The results showed that the proposed approach could not only reduce items drastically with the same classification precision, but also perform better on identifying different stages of AD comparing with other existing scales.

## 1. Introduction

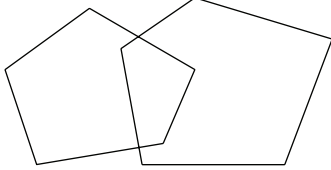
Alzheimer's disease (AD) is a degenerative senile dementia characterized by memory loss and cognitive functions disorders, and it is also one of the main types of senile dementia [1]. As AD has a slow onset and no highly specific diagnostic indicators at the early stage of the disease, it is particularly challenging for primary clinicians to identify transition points (from the asymptomatic phase to the symptomatic predementia phase to dementia onset) for individual patients [2, 3]. It is, nevertheless, important to identify these transition points between different stages, because studies [4] have proved that targeted therapies may help slow down the progress of the disease and improve quality of life for patients and their families.

Due to the lack of advanced medical facilities (advanced imaging and cerebrospinal fluid measures), the screening

of AD usually depends on the use of neuropsychological rating scales in primary clinics. Various neuropsychological rating scales, which are considered as a reliable and valid standardized testing tool, have been designed for cognitive abilities screening, and many of them have yielded good results as decision-making tools, such as minimal state examination (MMSE) [5], clinical dementia rating (CDR) [6], Montreal Cognitive Assessment (MoCA) [7], Geriatric Depression Scale (GDS) [8], and Activity of Daily Living Scale (ADL) [9]. Table 1 and Figure 1 are two most commonly used rating scales (the MMSE and the MoCA) in clinical practice.

However, each neuropsychological rating scale has its emphasis and limitation. A previous study has shown some scales do not perform well in one or more cognitive domains [10]. Multiple neuropsychological rating scales can cover more comprehensive cognitive domains. Therefore, multiple scales should be used together in order to get patients'

TABLE 1: Minimal state examination (MMSE).

Orientation	Year Month Day Date Time	___/5
	Country Town District Hospital Ward	___/5
Registration	Examiner names 3 objects (e.g. apple, table, and penny). Patient asked to repeat (1 point for each correct answer). THEN patient to learn the 3 names repeating until correct.	___/3
Attention and calculation	Subtract 7 from 100 and then repeat from result. Continue 5 times: 100 93 86 79 65 Alternative: spell "WORLD" backwards-"DLROW"	___/5
Recall	Ask for names of 3 objects learned earlier.	___/3
Language	Name a pencil and watch	___/2
	Repeat "No fits, ands, or buts"	___/1
	Give a 3-stage command. Score 1 for each stage. E.g., "Place index finger of right hand on your nose and then on your left ear"	___/3
	Ask patient to read and obey a written command on a piece of paper stating "Close your eyes"	___/1
	Ask patient to write a sentence. Score if it is sensible and has a subject and a verb	___/1
Copying	Ask the patient to copy a pair of intersecting pentagons:	
		___/1
Total		___/30

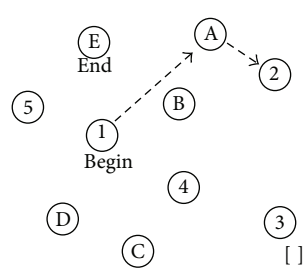
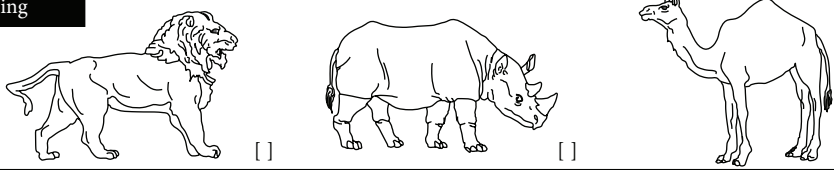
comprehensive cognitive status, which can help doctors to make correct diagnosis. However, this will bring two challenges: (1) neuropsychological testing requires highly trained assessors [11], while most primary clinicians are not qualified to conduct a full mental status examination or interpret a battery of scales' score; it is difficult for them to offer exact judgments about the examinee's cognitive state [12]. (2) Neuropsychological testing is quiet time consuming; the elders cooperate well only for short periods with the limitation of vitality and cognition [13], so long-time testing will bring negative impact on the quality of neuropsychological testing. Thus, we can conclude that the screening of AD in primary clinics should be based on the criteria that can get maximum accuracy in a convenient way within limited time.

To solve the above-mentioned challenges, identifying the items with the best ability to distinguishing AD (called critical items for short) from a battery of commonly used rating scales may help improve the efficiency of cognitive abilities screening. Then, a well-performance decision-making model, while the previously selected items can be taken as its input, may help primary clinicians improve diagnostic accuracy in routine clinical practice. So in this paper, we suggest dealing with the screening of AD by means of a two-stage hybrid intelligent approach based on multineuropsychological rating scales analysis: in Stage 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Stage 2, use a Bayesian network to develop a diagnosis assisting model of AD based on the selected items. This hybrid intelligent technique takes the advantage of attributes reduction of rough set theory requiring no prior knowledge and the uncertain reasoning ability of Bayesian network to build a relatively convenient and accurate decision-making model for primary clinicians.

The rest of this paper is organized as follows: Section 2 introduces the related work; Section 3 introduces basic concepts behind rough set theory, genetic algorithm, and Bayesian network; Section 4 presents the proposed approach including the proposed GA-RS attributes reduction algorithm applied in AD and the Bayesian network model constructed for AD diagnosis; Section 5 describes the evaluation results of the proposed model; then in Section 6, the discussions on some benefits and limitations of the proposed approach in the clinical environment are made; Section 7 draws conclusions and future work.

## 2. Related Work

Since the application of multiple neuropsychological rating scales is very time consuming and challenging for primary care physicians, many researchers have been trying to find out the most effective screening method for clinical application. For example, some scholars have tried to simplify the MMSE, which is the most widely used scale for screening dementia. Lou et al. [14] reported a 16-item simplified version of the original MMSE with high sensitivity and specificity. Callahan et al. [15] designed a six-item screener, which derived from the MMSE, the Blessed Dementia Rating Scale (BDRS), and the Word List Recall. And its sensitivity and specificity for a diagnosis of dementia reached 88.7% and 88.0%. In addition to the study of the MMSE, some researchers also have studied the combination of multiple neuropsychological rating scales. For example, Chen et al. [16] proposed an eight-item test, subtracted from the MMSE, the Clock Drawing Test (CDT), and the Instrumental Activities of Daily Living Scale (IADL). The evaluation result revealed that it was a sufficient and simple tool for the screening of early dementia in primary

Montreal cognitive assessment (MOCA)		Name:		Date of birth:			
Version 7.1 Original version		Education:		Sex:			
				Date:			
<b>Visuospatial/executive</b>		Copy cube	Draw CLOCK (ten past eleven) 3 points			<b>Points</b>	
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	___/5	
			Contour	Numbers	Hands		
<b>Naming</b>						___/3	
<b>Memory</b>	Read list of words, subject must repeat them. Do 2 trials, even if 1st trail is successful. Do a recall after 5 minutes	Face	Velvet	Church	Daisy	Red	No points
	1st trail						
	2nd trail						
<b>Attention</b>	Read list of digits (1 digit/s)	Subject has to repeat them in the forward order			<input type="checkbox"/>	2 1 8 5 4	___/2
		Subject has to repeat them in the backward order			<input type="checkbox"/>	7 4 2	
	Read list of letters. The subject must tap with his hand at each letter A. No points if >2 errors	<input type="checkbox"/> F B C M N A A J K L B A F A K D E A A A J A M O F A A B					___/1
	Serial 7 subtraction starting at 100	<input type="checkbox"/> 93	<input type="checkbox"/> 86	<input type="checkbox"/> 79	<input type="checkbox"/> 72	<input type="checkbox"/> 65	___/3
	4 or 5 correct subtractions: 3 pts, 2 or 3 correct: 2 pts, 1 correct: 1 pt, 0 correct: 0 pt						
<b>Language</b>	Repeat: lonely know that John is the one to help today. <input type="checkbox"/>					___/2	
	The cat always hid under the couch when dogs were in the room. <input type="checkbox"/>						
	Fluency/name maximum number of words in one minute that begin with the letter F. <input type="checkbox"/> ____(N ≥ 11 words)					___/1	
<b>Abstraction</b>	Similarity between e.g. banana-orange = fruit <input type="checkbox"/> train-bicycle <input type="checkbox"/> watch-ruler					___/2	
<b>Delayed recall</b>	Has to recall words WITH NO CUE	Face	Velvet	Church	Daisy	Red	Points for UNCUED recall only
		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Optional</b>	Category cue						
	Multiple choice cue						
<b>Orientation</b>	<input type="checkbox"/> Date	<input type="checkbox"/> Month	<input type="checkbox"/> Year	<input type="checkbox"/> Day	<input type="checkbox"/> Place	<input type="checkbox"/> City	___/6

Normal ≥ 26/30 Total \_\_\_/30  
Add 1 point if ≤12 year edu

FIGURE 1: Montreal Cognitive Assessment (MOCA).

care clinics. Besides, other studies [17–20] have reported various screening methods such as instruments designed for detection of memory, attention-executive function, visuospatial ability, and interview with reliable informants. After a general survey of the above studies, we found that all

these researchers selected items according to their own clinical intuition regarding domains of impairment commonly encountered in AD rather than objective data analysis based on past cases. This selection was influenced by subjective factor easily. After selecting discriminative items, a large

number of experiments need to be done for confirming weight coefficient of each item. For instance, the MoCA team spent over 5 years modifying the MoCA for clinical use [7]. So a method based on data mining may bring a new thought to searching the most discriminative items from multiple neuropsychological rating scales. It is based on large-scale objective clinical evidences, and, moreover, it can not only improve efficiency of identifying critical items but also adjust weight coefficient of items automatically.

The computer-aided diagnosis of AD is always the hot topic of research in the last few years. DeFigueiredo et al. [21] presented an algorithm based on the analysis of computer tomography image (CT) data from brain. This algorithm used an optimal interpolative neural network to classify individuals into four different groups (i.e., clinically diagnosed groups of elderly normal, demented, AD, and vascular dementia subjects). Ramírez et al. [22] showed a computer-aided diagnosis (CAD) system for the diagnosis of AD. His method is based on partial least squares regression model and a random forest predictor. The analyzed data is from single photon emission computed tomography (SPECT). Further works on SPECT data have reported high AD classification accuracy [23–25]. Duchesne et al. [26] presented their work on automated computer classification in Alzheimer’s dementia using the context of cross-sectional analysis of magnetic resonance images (MRI). Daliri [27] also presented an automated method for diagnosing AD from brain MR images. In addition to volumetric MRI, diffusion tensor imaging (DTI) has increasingly been used to detect microstructural brain differences in AD. Graña et al. [28] obtained discriminant features from two scalar measures of DTI data and used support vector machine (SVM) as a classifier. From the view point of methodology, most of the studies focused on finding the microstructural differences using image analysis based on supervised machine learning algorithms. Although the analysis based on imaging is becoming an important research trend of AD studies, the technique is difficult to be applied in primary clinical settings because access to imaging equipment may be limited there.

Based on above-mentioned analysis, we proposed a computer-aided quick screening method for AD based on multiple neuropsychological rating scales. We especially used data mining technique to reduce the items of scales and lower the barriers to applying this method in primary clinical settings.

### 3. Preliminary

This section addresses some basic concepts needed for the remainder of the discussion. We introduce rough set theory first and then discuss genetic algorithm and Bayesian network, so as to set up a necessary context for describing our approach.

*3.1. Rough Set Theory.* Rough set theory [29–31] is a mathematical approach to deal with imprecision, vagueness, and uncertainty. The main difference between rough set theory and other mathematical tools for dealing with uncertain

problems is that rough set theory does not need any prior information beyond the problem itself. Rough set theory is an important method for attribute reduction, where attributes that do not contribute to the classification of the given training data can be identified and removed.

Rough set theory is based on the establishment of equivalence classes within the given training data. All the data tuples forming an equivalence class are indiscernible; that is, the samples are identical with respect to the attributes describing the data. Formally, the indiscernibility relation is defined as follows:

$$\text{IND}(B) = \{(x, y) \in O^2 \mid \forall a \in B, f(x, a) = f(y, a)\}, \quad (1)$$

where  $A$  is a nonempty finite set of attributes and  $O$  is a nonempty finite set of objects. Given any  $B \subseteq A$ , relation  $\text{IND}(B)$  induces a partition of  $O$ , which is denoted by  $O/\text{IND}(B)$ , where an element from  $O/\text{IND}(B)$  is called an equivalence class.

It is common that some classes cannot be distinguished by given real-world data in terms of the available attributes. Rough sets can be used to approximately or “roughly” define such classes. A rough set definition for a given class  $X$ ,  $X \subseteq O$ , is approximately by two sets—a lower approximation of  $X$  and upper approximation of  $X$ . The lower approximation of  $X$  consists of all the data tuples that, based on the knowledge of the attributes, are certain to belong to  $X$  without ambiguity. The upper approximation of  $X$  consists of all the tuples that, based on the knowledge of the attributes, cannot be described as not belonging to  $X$ . So the lower and upper approximation of  $X$ , is defined, respectively, as

$$\underline{X}(B) = \{x \in U \mid [x]_B \subseteq X\}, \quad (2)$$

$$\overline{X}(B) = \{x \in U \mid [x]_B \cap X \neq \Phi\},$$

where  $x \in O$  and the pair  $(\underline{X}(B), \overline{X}(B))$  is called the rough set with respect to  $X$ . The set  $BN_B(x) = \overline{X}(B) - \underline{X}(B)$  is called the boundary region of  $X$ . Let  $[x]_B$  denote the equivalence class of relation  $\text{IND}(B)$  that contains element  $x$ .

An important concept in attributes reduction is dependency of attributes, which can be defined in the following way.

For  $A$  and  $B$ ,  $B$  depends on  $A$  in degree  $\gamma_B$ ,

$$\gamma_B = \frac{|\text{POS}_A(B)|}{|O|}, \quad (3)$$

$$\text{POS}_A(B) = \bigcup_{X \in O/\text{IND}(B)} \underline{X}_A, \quad (4)$$

where  $\gamma_B = 1$  means  $B$  depends totally on  $A$ .  $0 < \gamma_B < 1$  indicates  $B$  depends partially on  $A$ , and if  $\gamma_B = 0$ ,  $B$  is totally independent on  $A$ .

Formally, a reduct is a subset of attributes which can fully characterize the knowledge in the database. Let  $\text{RED}(B)$  be the minimal subset of  $A$  and  $\text{CORE}(B)$  the set of attributes which cannot be eliminated and the intersection of all reducts. If the attribute in  $\text{CORE}(B)$  is removed, the ability to classify objects into the elementary classes of  $B$  will decrease:

$$\text{CORE}(B) = \bigcap_{R_i \in \text{RED}(B)} R_i, \quad i = 1, 2, 3, \dots \quad (5)$$

**3.2. Genetic Algorithm.** The genetic algorithm (GA) [32–34] is an optimized algorithm based on the Darwinian principle of natural selection. It can be used with other data mining techniques for optimization and performance amelioration. The genetic algorithm process starts with the randomly generated and encoded initial population, which includes several hundreds or thousands of potential solutions to the problem. Each encoded individual in the population is called chromosome and each bit in the chromosome is called gene and has a value. The next step is called genetic operators. The most widespread genetic operators include selection, crossover, and mutation. Each chromosome in the population is evaluated by user-defined fitness function. The higher a chromosome's fitness value is, the more likely it is to produce offspring. In this way the overall fitness of the population is guaranteed to increase and those with weak fitness will be eliminated gradually. Crossover forms new chromosomes for the population by exchanging a fixed part between two chromosomes. The chromosomes most often used for crossover are those destined to be eliminated from the population. Mutation can be applied by randomly flipping bits (or attribute values) within a single chromosome to avoid the local optima. New offspring is reevaluated by fitness function to search the solution. The whole process is repeated until reaching the prespecified number of generations or the desired level of fitness.

**3.3. Bayesian Network.** Bayesian network is an acyclic directed graph for representing probabilistic relationships among a set of random variables [35]. Trained Bayesian networks can be used for classification. There are two key elements of a Bayesian network: (1) a directed acyclic graph (DAG) encoding the dependence relationships among a set of variables and (2) a probability table associating each node to its immediate parent nodes. Each node in the directed acyclic graph represents actual attributes given in the data. Each arc represents a probabilistic dependence. If an arc is drawn from a node  $X$  to a node  $Y$ , then  $X$  is a parent of  $Y$ , and  $Y$  is a descendant of  $X$  [36]. Bayesian network has one conditional probability table (CPT) for each node. The CPT for a node  $Y$  specifies the conditional distribution  $P(Y \mid \text{Parents}(Y))$ , where  $\text{Parents}(Y)$  are the parents of  $Y$ . A node within the network can be selected as an “output” node, representing a class label attribute. There may be more than one output node. Given a set of variables, the network can be used to compute the probabilities of the presence of various classes, rather than return a single class label. There have been some works with applications using Bayesian network in diagnosis of AD [37, 38].

## 4. Methods

**4.1. Overview.** In this section, we present the formation process of AD diagnosis assisting model with the proposed hybrid intelligent method, which consists of two steps: in Step 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Step 2, use a Bayesian network to build a diagnosis assisting model of AD based on selected items.

In the first step, finding critical items from multiple neuropsychological scales is the problem of attributes reduction, which is also a classical problem in machine learning. Rough set theory is a useful attributes reduction method in machine learning. It can find the shortest or minimal reducts while keeping high-quality classification performance [39]. However, current rough set approaches to attributes reduction are inadequate to find optimal reductions as no perfect heuristic can guarantee optimality. Optimal attribute reduction has been proved to be a NP-hard problem [40]. So, stochastic approaches provide a promising attributes reduction mechanism, like genetic algorithm (GA). GA is an effective and robust method for solving both constrained and unconstrained multiparameter optimization problems that is based on natural selection. Many literatures have combined rough set theory and genetic algorithm for solving machine learning problems in a variety of domains [41–44]. We present a genetic algorithm for reduct set computation which is very fast and gives a good approximation in the AD field.

As exact causes and the mechanism of AD remain uncertain, there is currently no method to ensure the AD presence except for an autopsy. Clinicians can only make the diagnosis called “probable or possible AD” in clinical environment, especially in the early stage of this disease. The proposed approach applies Bayesian network that has strong reasoning ability in solving uncertain problems to build the decision-making model and predict the AD probability rather than offer a definitive diagnosis. So it is more conducive to the practical application of the proposed approach.

**4.2. Attributes Reduction Based on Genetic Algorithm and Rough Set Theory.** As mentioned above, GA-RS is used to identify critical items from a battery of rating scales. Each step of the algorithm is described as follows.

**4.2.1. Chromosome Representation.** Because genetic algorithm cannot deal with data in solution space directly, we must represent them as binary strings of length  $M$  which is the number of the condition attributes by encoding. Binary encoding is simple and easy to operate. Each binary string is called a chromosome, in which “1” means that the corresponding attribute is selected and “0” means not. Attributes in *Core* should take “1”, and remain the same in the whole process of evolution, since genetic search starts from the *Core*.

**4.2.2. Fitness Function.** Fitness function is a user-defined function which is used to measure each chromosome's optimization calculation in the groups. The fitness value of each chromosome represents suitability for the environment. In this paper, we expect the “best” chromosome could have the minimal length and the strongest classification performance as the algorithm proceeds. So the fitness function is defined as follows:

$$\begin{aligned}
 F(x) &= \beta \cdot f(x) + p(x) \\
 &= \beta \cdot \left( 1 - \frac{\text{card}(x)}{\text{card}(C)} \right) + \frac{\text{card}(\text{POS}_x(D))}{\text{card}(\text{POS}_C(D))}, \quad (6)
 \end{aligned}$$

where  $\text{card}(x)$  is the number of “1” in chromosome, which means the number of condition attributes contained by chromosome;  $\text{card}(C)$  is the length of chromosome, which is the total number of condition attributes;  $f(x) = 1 - \text{card}(x)/\text{card}(C)$  indicates the chromosome  $x$  that is not included in the proportion of condition attributes.  $p(x)$  indicates the distinction ability of attribute  $x$ .

**4.2.3. Selection Method.** Select chromosomes based on their fitness values from the current population to produce offspring for the new population. Tournament selection is used, which means the higher the fitness value is, the higher probability of that chromosome is selected for reproduction. This step is repeated until the number of chromosomes selected is equal to the number of the population.

**4.2.4. Crossover and Mutation.** One-point crossover method is used to reproduce with a probability of  $P_c$ . In mutation process, we first select a chromosome to be mutated with probability  $P_m$  and then replace a single gene of the chromosome from “1” to “0” or from “0” to “1” randomly.

**4.2.5. Elitist Strategy.** We take the elite strategy [45] to preserve the best individual of the fitness function value. Copy the individual of highest fitness value in the current generation to the next generation, unaltered.

The detail of the whole algorithm is as follows.

**Input.** Decision table  $IS = \langle O, A, V, f \rangle$ ;  $O$  is a nonempty finite set of objects.  $A$  is a nonempty finite set of attributes:  $A = C \cup D$ ,  $C$  is the set of condition attributes, and  $D$  is the set of decision attributes.  $V = \bigcup_{a \in A} V_a$ ;  $V_a$  is the set of values of attribute  $a \in A$ .  $f : O \times A \rightarrow V$  is an information function so that, for any  $a \in A$  and  $x \in O$ ,  $f(x, a) \in V_a$ .

**Output.** There is an attributes reduction  $R$  of decision table.

#### Steps

**Step 1.** Calculate the dependency  $\gamma_c(D)$  between decision attributes set  $D$  and condition attributes set  $C$  by formula (3).

**Step 2.** Let  $\text{Core}(C) = \varphi$ , to get rid of each attribute  $c \in C$  one by one, if  $\gamma_{C-c} \neq \gamma_C$ ,  $\text{Core}(C) = \text{Core}(C) \cup \{c\}$  which means the core is  $\text{Core}(C)$ ; if  $\gamma_{\text{Core}(C)}(D) = \gamma_C(D)$ , then the core is minimal attributes reduction and if not, go to Step 3.

**Step 3.** Generate  $m$  binary strings with length  $n$  randomly, which can be seen as the initial population.  $n$  is the number of the condition attributes. “1” means that the corresponding attribute is present, and “0” indicates not. For attributes in core, corresponding position is “1” and for others, corresponding position is “1” or “0” randomly.

**Step 4.** Calculate the fitness value for each individual by formula (6) and select individuals by tournament selection.

**Step 5.** Perform crossover operation according to the crossover probability  $P_c$ , using single-point crossover mode.

TABLE 2: Parameter settings of GA.

Population size	1000
Number of generations	500
Initialization method	Binary method
Percentage of elite	0.2
Selection method	Tournament selection
Crossover method	Uniform crossover
Crossover ratio	0.5
Mutation method	Single-point mutation
Mutation ratio	0.03

**Step 6.** Perform mutation operation according to the mutation probability  $P_m$ . We basically bit mutation strategy while the corresponding bit of attributes in the *Core* does not change.

**Step 7.** Select the individuals with the best fitness values to be offspring of the current generation. This strategy is to guarantee the best chromosome could carry over to the next generation.

**Step 8.** Repeat the genetic operation until either one of the following conditions is satisfied: (1) the maximum number of generations is achieved or (2) the fitness value of the best individual for the present generation no longer changes during several successive generations.

**Step 9.** Convert the best individual to condition attribute and get the final result.

The whole computation steps are shown in Figure 2.

Parametric settings of genetic algorithm are as follows: population scale  $N = 1000$ , crossover ratio  $P_c = 0.5$ , mutation ratio  $P_m = 0.03$ , and the largest number of iterations is 500, just as demonstrated in Table 2.

The fitness function employed in this paper controls the chromosomes that evolve in the direction of the minimum reduction while keeping the classification performance: the higher the  $\text{card}(x)$  is, the smaller the  $f(x)$  is; the larger  $p(x)$ , the more dependence between the condition attribute  $C$  and decision attribute  $D$ . This algorithm ensures the two requirements, so the result is the optimal solution of the problem.

In our approach, attributes reduction mentioned above is not the final goal but an intermediate process and core technology of AD diagnosis assisting for clinician in primary clinic. An uncertainty inference model for AD should be built after attributes reduction, which will be discussed in next section.

**4.3. Bayesian Network Model for AD Diagnosis.** Based on the above step, we attempt to construct the structural model for AD diagnosis. These selected items can be represented as input variables of the model. Since there is strong diagnostic uncertainty earlier in the disease process, an uncertainty inference model must be built. A popular modeling tool for complex uncertain domains is a Bayesian network.

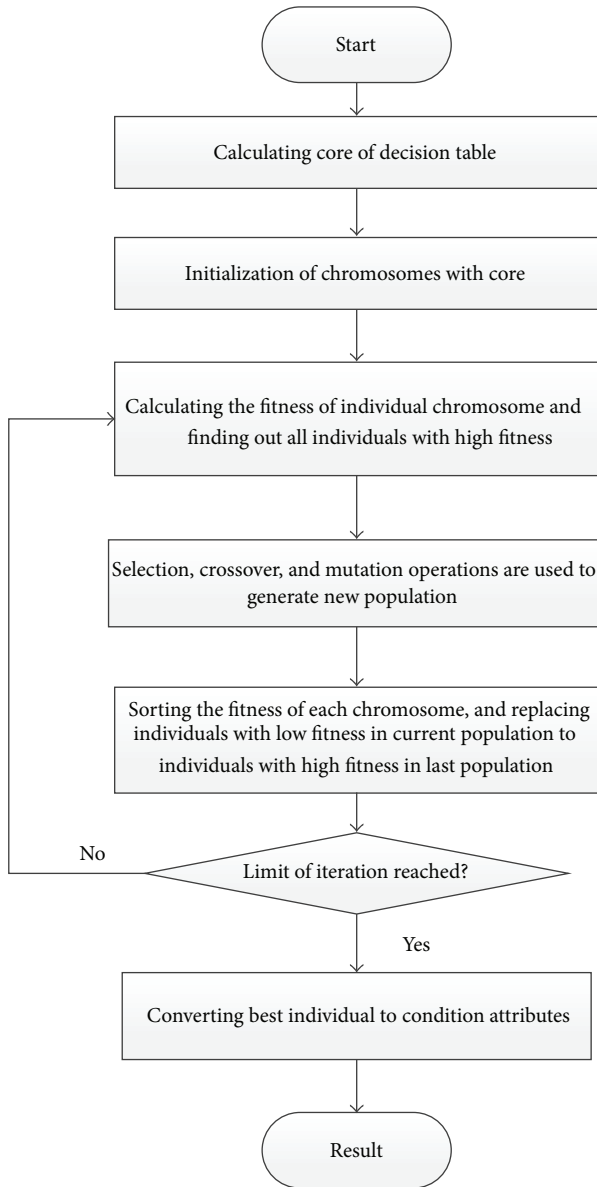


FIGURE 2: Detailed algorithm flow of GA-RS.

## 5. Experiment and Results

**5.1. Data Collection.** The experimental data set is composed of 500 consecutive historical cases collected by the neurology department of a certain top hospital in China from 2009 to 2014. Each case is a series of scale scores belonging to one subject, and each subject has only one case. All neuropsychological tests were conducted by trained neuropsychologists and administered on the same day. The mean age of subjects is 74.4 (range, 51–92); 59.5 of the subjects' percent were female. These 500 historical cases have the following characteristics.

- (1) All these 11 neuropsychological rating scales are selected from a large number of scales by leading experts in neurology, including the MMSE, the

MoCA, the CDR, the GDS, the ADL, the Word-List Learning, the figure copying, the new word discriminating, the trail making test, the similarity, and the perception. All these neuropsychological rating scales are commonly used instruments for screening cognitive or noncognitive impairment in the clinical diagnosis.

- (2) Each scale consists of a series of items. In total, there are 101 testing items in these scales. Some items are straightforward Q & A pattern, for instance, "What is the date?" Some others need the subject to do some actions, "Please read this and do what it says. (Show subjects the following words on the stimulus form: Close your eyes.)" Each of the tests scores points if it is answered correctly.

To ensure the correctness of diagnosis of each case, an expert panel group composed of three neuropsychologists was set up, and the diagnosis of each case was determined by the panel. The diagnosis of experts not only depended on an objective neuropsychological testing, but also on the history-taking from the patient and a knowledgeable informant. Their diagnosis was regarded as the gold standard. In the current study, the diagnosis of cases could be divided into three types: patients with AD, patients meeting criteria for mild cognitive impairment [46] (called MCI for short, which is regarded as the predementia stage of AD), and the elderly subjects with normal cognition, in which, the number of each type is 33.5%, 37.7%, and 28.8%, respectively.

Parts of cases are given in Table 3. In the table, each column is one testing item of scales, for instance, Time Orientation, Place Orientation and Repetition belong to the MMSE while Visuospatial Skills belongs to the MoCA. They are regarded as the condition attributes. The last column, Result, is the decision attribute (the diagnosis of each patient).

**5.2. Experimental Design and Results.** To verify the feasibility and validity of the proposed approach, the performance of proposed approach can be measured by the following evaluations: (1) reduction ratio on testing duration and reduction ratio on quantity of items; (2) comparison with multiple classifiers; (3) comparison of classification accuracy before and after reduction; (4) the performance of classification compared with two existing cognitive screening scales.

We applied two evaluation methods to prove the reliability of our experimental results, one was 10-fold cross-validation, and the other was 0.632 bootstrap. Obtain computing results by averaging after executing 10 times. Recall rate, precision rate, and accuracy were selected as the performance evaluation metrics. The obtained results of the experiment are to be presented and discussed in the next section.

**5.3. Reduction Result.** After attributes reduction, 10 items were selected finally, which are listed in Table 4.

Some items represent one test, such as Figure\_Copy, Figure\_Short\_Memory, Figure\_Delay\_Memory, Word\_Delay\_Recall, and Reading\_Comprehension, while others are a set of several tests, which cannot be separated (because these

TABLE 3: Part of AD dataset.

Fact	Time orientation	Place orientation	Repetition	Visuospatial skills	...	Result
1	5	5	1	5	...	Normal
2	4	5	1	5	...	Normal
3	4	5	1	3	...	MCI
4	1	4	0	3	...	AD
⋮	⋮	⋮	⋮	⋮	⋮	⋮
500	0	0	1	1	...	AD

TABLE 4: Reduction results by GA-RS.

Result	Source
Reading_Comprehension	MMSE
Visuospatial_Execution	MoCA
Naming	MoCA
Attention	MoCA
Figure_Copy	Figure copying test
Figure_Short_Memory	Figure copying test
Figure_Delay_Memory	Figure copying test
IADL	ADL
Word_Delay_Recall	Word-List Learning test
Word_AVG	Word-List Learning test

tests significantly correlate with one another and must be performed together), such as Visuospatial\_Execution, IADL, Naming, Attention, and Word\_AVG.

**5.3.1. Reduction Ratio.** We used reduction ratio including the reduction ratio of testing duration and the reduction ratio of quantity of items as measurable metrics. Assume that the number of condition attributes before and after reduction is  $m$  and  $n$ , respectively. The reduction ratio  $r$  is defined:

$$r = \frac{(m - n)}{m} \times 100\%. \quad (7)$$

Before reduction, the number of items is 34, while only 10 items left after reduction using the proposed method, so we can conclude that the reduction ratio is 70.59%. The experimental results indicate that using GA-RS to select subset can reduce items dramatically.

Similarly, the reduction ratio of testing duration can also be calculated using formula (7). In clinical practice, the duration of finishing these scales varies a lot, which depends on the subject's state of cognitive impairment. According to [47, 48], the performance time for the MMSE and the MoCA is 13.4 minutes and 14.8 minutes on average, respectively. Based on the past experience, a skilled clinician administers the scale for more than one hour to complete all the 11 scales mentioned above. By using the proposed model, clinicians do not have to finish all the scales but only need to complete the selected testing items. Hence the test duration is reduced greatly and ranges from 12 to 15 minutes with a mean time of 13.5 minutes and a standard deviation of 2.3 minutes.

**5.3.2. Comparison with Multiple Classifiers.** The constructed Bayesian network structure is presented in Figure 3.

We compared some common used classifiers with Bayesian network in order to select the well-performed classifier. All these classifiers had the same input items (the items selected by the above step). The result of comparison is as shown in Table 5.

From Table 5, we could see that the Bayesian network performed best in the four classifiers. Then, in order to further prove the effectiveness of results, we compared the four groups using Friedman test to see if a significant difference emerged (Table 6). Table 6 shows the mean rank for each classifier and Table 7 shows the result of the Friedman test.

In Table 7,  $P = 0.001 < 0.05$ , there was a significant difference between these four classifiers. However, the cross-validation estimate of prediction error may lead to a high variability in results. In order to validate the result further, 0.632 bootstrap was used to evaluate the performance, as shown in Table 8.

Similar result came from Table 8 when compared with Table 5. It suggested that Bayesian network performed better than other three classifiers. And Friedman test was also performed, as shown in Tables 9 and 10. We found that there existed a significant difference between the four groups as well. Because of low variance with only moderate bias, the result got by 0.632 bootstrap was selected as the final result.

**5.3.3. Classification Performance before and after Reduction.**

In order to evaluate the validity of attributes reduction, we used Bayesian network algorithm to compute the classification performance before and after attributes reduction, respectively, and to check whether or not the classification performance had changed.

Each subject had been given the probability of each classification. The highest probability was regarded as the diagnosis of the model. Table 11 presents a summary of the classification results before and after reduction.

From the variance of recall rate and precision rate after attributes reduction as shown in Table 11, we found that the recall rate and precision rate of each group decreased a little, but less than 3.05%. We analyzed the result data using Wilcoxon Signed-Rank Test. The calculated  $P$  value was 0.853 and larger than 0.05, so the null hypothesis was true, which means that there was no significant statistical difference between these two methods. In conclusion, the comparative experimental results indicated that the proposed



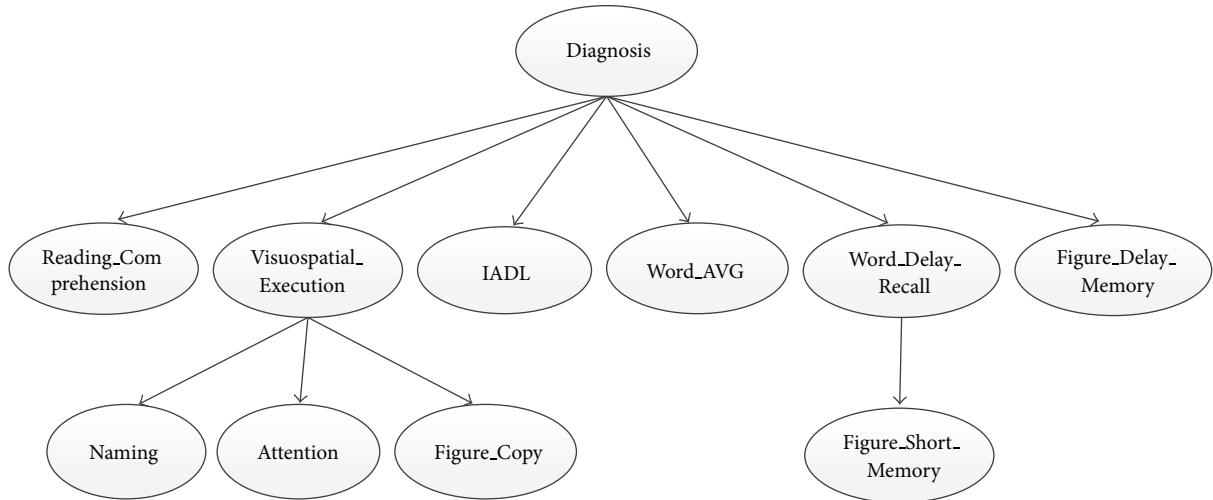


FIGURE 3: AD-related diagnosis assisting model.

TABLE 5: The comparison of classifiers by 10-fold validation.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
BN	85.41	87.23	78.95	73.17	90.7	95.12	85.27
NN	77.27	79.07	69.05	64.44	86.05	90.24	77.52
C4.5	81.82	63.16	42.86	48.65	72.09	88.57	65.89
SMO	86.36	74.51	64.29	77.14	93.02	93.02	81.35

R: recall rate; P: precision rate; ACC: accuracy; BN: Bayesian network; NN: neural network.

TABLE 6: Ranks in Friedman tests.

	Mean rank
Bayesian network	2.58
Decision tree (C4.5)	2.27
Neural network	2.50
SMO	2.66

TABLE 7: Test statistics in Friedman tests.

Chi-square	15.789
df	3
Asymp.Sig.	0.001

method could find the shortest or minimal reducts while keeping high-quality classification performance.

5.3.4. *Comparison with Comprehensive Cognitive Screening Scales.* Comprehensive cognitive screening scales measure all important aspects of cognitive function, such as memory, language, visuospatial skills, attention, and executive function. The most commonly used comprehensive cognitive screening scales include the MMSE and the MoCA. Our computer-aided model also covers multiple aspects of cognitive screening, so the comparison of recall rate and precision rate between our model and these two scales is needed based on the same dataset in order to evaluate the validity of our model.

TABLE 8: The comparison of classifier by 0.632 bootstrap.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
BN	84.55	84.91	75	76.39	91.94	89.81	83.73
NN	78.99	71.76	60.63	67.68	91.55	91.98	76.94
C4.5	71.90	66.67	46.45	50.26	79.15	80.29	66.24
SMO	84.71	79.46	63.03	69.63	89.57	87.91	79.61

R: recall rate; P: precision rate; ACC: accuracy; BN: Bayesian network; NN: neural network.

TABLE 9: Ranks in Friedman tests.

	Mean rank
Bayesian network	2.59
C4.5	2.30
Neural network	2.54
SMO	2.57

TABLE 10: Test statistics in Friedman tests.

Chi-square	48.694
df	3
Asymp.Sig.	0.000

TABLE 11: Comparison of the classification performance before and after attributes reduction by Bayesian network.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
Before	86.36	84.44	76.19	76.19	90.7	92.86	84.50
After	84.55	84.91	75	76.39	91.94	89.81	83.73

R: recall rate; P: precision rate; ACC: accuracy.

The MMSE is a questionnaire test that is used to screen cognitive impairment. The total score is 30. If the score is greater than or equal to 27 points, it means the subject has

TABLE 12: Comparison with the MMSE and the MoCA.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
MMSE	93.88	67.65	45.83	61.11	84.29	93.65	74.29
MoCA	68.89	86.11	81.82	60.00	80.39	93.18	77.14
Model	84.55	84.91	75.00	76.39	91.94	89.81	83.57

R: recall rate; P: precision rate; ACC: accuracy.

a normal cognition. Below 27 points, scores can be divided into several stages: severe cognitive impairment ( $\leq 9$  points), moderate cognitive impairment (10–18 points), and mild (19–24 points) cognitive impairment [49]. The MoCA is also a one-page 30-point test developed as a brief cognitive screening tool to detect mild-moderate cognitive impairment. The suggested cut-off score on MoCA is 26, which yielded the best balance between sensitivity and specificity for the MCI and AD groups. We applied these criteria to our dataset, and the result is showed in Table 12.

From Table 12, we found that the MMSE did not perform well as a screening instrument for MCI due to the lack of sensitivity to MCI [50]. Some researchers believe the low sensitivity of the MMSE comes from the emphasis placed on language items and a paucity of visuospatial items [51]. However, visuospatial skills and executive function had been retained through our attributes reduction algorithm. So our model performed much better than the MMSE on detecting MCI, which is a very significant stage of AD. In general, our model is a more effective tool for identifying different stages of AD than the MMSE.

The MoCA is designed for the detection of MCI; that is to say, it is developed to screen patients who has cognitive impairment complaints but still performed in the normal range on the MMSE. The MoCA is more sensitive on detecting MCI than the MMSE, because the MoCA focuses more on tasks of frontal executive functioning and attention. Our model retained these key parts of the MoCA, so its performance on detecting MCI was close to that of the MoCA. Above all, our model was advantageous when identifying multiple transition points between different stages of AD, and it was not only designed for screening MCI. Compared with the MoCA, our model had distinct advance on differentiating normal and AD while almost keeping the sensitivity of detecting MCI. This was more helpful for primary clinicians to take target care and therapies.

We also performed Friedman test on these three groups and the actual result of the Friedman test is shown in Tables 13 and 14. From the result, we can see that there is an overall statistically significant difference between the mean ranks of these three groups.

## 6. Discussion

This study proposed a computer-aided diagnosis model of AD applied in primary clinics. In order to solve especially the problem that cognitive screening based on multiple neuropsychological scales was time consuming, GA-RS algorithm was used to identify the most related items from numerous rating scales while ensuring a satisfactory accuracy

TABLE 13: Ranks in Friedman tests.

	Mean rank
MMSE	1.94
MoCA	1.98
Model	2.08

TABLE 14: Test statistics in Friedman tests.

Chi-square	3.756
df	2
Asymp.Sig.	0.043

of classification. The experimental results validated the effectiveness of the proposed approach.

The benefits of this study are listed as follows.

- (i) The proposed approach is suitable to be applied in the primary clinics, because the clinician's day is labor-intensive, and the mean time of clinical interviewing is limited for each patient. Clinicians are eager to have a tool for the screening of AD without spending too much time. The proposed approach reduces the testing time for each patient while keeping classification accuracy. Hence, such computer-aided approach is applicable in clinical practice.
- (ii) As an important merit of the proposed approach, it is a computer-aided diagnostic tool based on multiple neuropsychological rating scales, rather than neuroimaging, or biomarker. To the best of our knowledge, there exist few reports providing a computer-aided diagnosis method that is completely based on multiple neuropsychological rating scales. Thus our approach is suitable to be popularized in the primary clinics which have no advanced imaging and biological molecular equipment.
- (iii) It is estimated by the specialists that there are 68 relevant AD scales in the world [52], most of which have established the normative data and interpretation of scores in different countries. However, there have been seldom studies on how to employ so many rating scales to give a comprehensive diagnosis. For instance, if a patient "MMSE = 27", "MoCA = 23", and "ADL = 26," then what is the comprehensive status of the patient? It is relatively difficult for young general practitioners to make diagnosis in primary clinics. Our approach provides a new thought to solve this problem in hopes of supplementing the research in this field.
- (iv) The proposed approach is based on neuropsychological rating scales. Any disease that has no specific golden criteria and needs a long test by rating scales can try this method.

It should also be mentioned that there remain some limits to the approach proposed in this paper.

First of all, the chosen data might be bias, as all the cases of the study were collected only in one hospital. Secondly, there are some other types of dementia, such as Lewy body

dementia and vascular dementia, which are difficult to be distinguished from AD for young practitioners. The cases of these diseases are not included in our dataset, so clinicians must differentiate diagnostic methods for these diseases when using our model. Thirdly, the classification performance of Bayesian Network does not perform as well as expected, more machine learning algorithm can be tried to improve the classification performance. Finally, the number of subjects is still limited and more subjects are necessary in order to generalize the results to a larger population.

## 7. Conclusion and Future Work

The increasing aging population has led to a high increase in the prevalence of AD. Due to the fact that targeted care and therapies may slow down the progression of disease, the identification of different stages of AD is very important. In this paper, we proposed a computer-aided diagnosis method for AD based on analyzing the practical scores of rating scales. We especially identified the most discriminative items based on rough set theory and genetic algorithm. The selected items cover multiple cognitive domains and can be administered generally within 15 minutes. So it is user-friendly and is quickly administered, it may be appropriate use in primary clinics where assessment time is often limited. By comparing the classification performance, the result showed that the approach can effectively reduce the representation space of the attributes whilst hardly decreasing classification precision. The data also indicated that it has satisfactory reliability for both MCI and AD comparing with other existed cognitive screening scales.

Without doubt, opportunities for future research are abundant. First, we plan to further evaluate the built model with a perspective study in a real clinical setting. Second, more rating scales for specific dementias are going to be involved in the training set data and more comprehensive model for senile dementia will be built in the future work. Based on above work, a “three-level medical service network” for AD is going to be built in the near future and different computer-aided diagnosis tools for each level hospital will be developed; for example, the simple cognitive screening tool helps clinicians in primary clinics to judge whether patients suffer from cognitive impairment; the advanced cognitive assessment tool helps clinicians in second class hospitals to estimate the severity of cognitive impairment; the comprehensive assisted diagnosis tool is designed for clinicians in top hospitals to differentiate the types of dementia. The setup of such network will improve diagnosis accuracy of AD greatly and reduce the burden on public health care resource.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the National High Technology Research and Development Program of China (863 Program)

under Grant no. 2012AA02A601 and the National Science and Technology Major Project of China under Grant no. 2013ZX03005012.

## References

- [1] C. S. Porto, H. C. Fichman, P. Caramelli, V. S. Bahia, and R. Nitrini, “Brazilian version of the Mattis dementia rating scale: diagnosis of mild dementia in Alzheimer’s disease,” *Arquivos de Neuro-Psiquiatria*, vol. 61, no. 2B, pp. 339–345, 2003.
- [2] R. A. Sperling, P. S. Aisen, L. A. Beckett et al., “Toward defining the preclinical stages of Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s & Dementia*, vol. 7, no. 3, pp. 280–292, 2011.
- [3] G. M. McKhann, D. S. Knopman, H. Chertkow et al., “The diagnosis of dementia due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s and Dementia*, vol. 7, no. 3, pp. 263–269, 2011.
- [4] J. Rhodes-Kropf, “Palliative care for patients with Alzheimer’s dementia: advance care planning across transition points,” in *Choices in Palliative Care*, pp. 144–156, Springer, 2007.
- [5] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ““Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician,” *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.
- [6] C. P. Hughes, L. Berg, and W. L. Danziger, “A new clinical scale for the staging of dementia,” *The British Journal of Psychiatry*, vol. 140, no. 6, pp. 566–572, 1982.
- [7] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., “The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment,” *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [8] J. A. Yesavage, T. L. Brink, T. L. Rose et al., “Development and validation of a geriatric depression screening scale: a preliminary report,” *Journal of Psychiatric Research*, vol. 17, no. 1, pp. 37–49, 1982.
- [9] P. Self-Maintenance, *Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living*, 1969.
- [10] T. N. Tombaugh and N. J. McIntyre, “The mini-mental state examination: a comprehensive review,” *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, 1992.
- [11] C. Zadikoff, S. H. Fox, D. F. Tang-Wai et al., “A comparison of the mini mental state exam to the Montreal cognitive assessment in identifying cognitive deficits in Parkinson’s disease,” *Movement Disorders*, vol. 23, no. 2, pp. 297–299, 2008.
- [12] S. Borson, J. Scanlan, J. Hummel, K. Gibbs, M. Lessig, and E. Zuhr, “Implementing routine cognitive screening of older adults in primary care: process and impact on physician behavior,” *Journal of General Internal Medicine*, vol. 22, no. 6, pp. 811–817, 2007.
- [13] H. Halstead, “A psychometric study of senility,” *The British Journal of Psychiatry*, vol. 89, no. 376–377, pp. 363–373, 1943.
- [14] M.-F. Lou, Y.-T. Dai, G.-S. Huang, and P.-J. Yu, “Identifying the most efficient items from the Mini-Mental State Examination for cognitive function assessment in older Taiwanese patients,” *Journal of Clinical Nursing*, vol. 16, no. 3, pp. 502–508, 2007.
- [15] C. M. Callahan, F. W. Unverzagt, S. L. Hui, A. J. Perkins, and H. C. Hendrie, “Six-item screener to identify cognitive impairment

- among potential subjects for clinical research,” *Medical Care*, vol. 40, no. 9, pp. 771–781, 2002.
- [16] C.-Y. Chen, K.-K. Leung, and C.-Y. Chen, “A quick dementia screening tool for primary care physicians,” *Archives of Gerontology and Geriatrics*, vol. 53, no. 1, pp. 100–103, 2011.
- [17] R. B. Lipton, M. J. Katz, G. Kuslansky et al., “Screening for dementia by telephone using the memory impairment screen,” *Journal of the American Geriatrics Society*, vol. 51, no. 10, pp. 1382–1390, 2003.
- [18] A. F. Jorm, “The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review,” *International Psychogeriatrics*, vol. 16, no. 3, pp. 275–293, 2004.
- [19] S. Borson, J. M. Scanlan, J. Watanabe, S.-P. Tu, and M. Lessig, “Simplifying detection of cognitive impairment: comparison of the Mini-Cog and Mini-Mental State examination in a multiethnic sample,” *Journal of the American Geriatrics Society*, vol. 53, no. 5, pp. 871–874, 2005.
- [20] S. Kilada, A. Gamaldo, E. A. Grant, A. Moghekar, J. C. Morris, and R. J. O’Brien, “Brief screening tests for the diagnosis of dementia: comparison with the Mini-Mental State Exam,” *Alzheimer Disease and Associated Disorders*, vol. 19, no. 1, pp. 8–16, 2005.
- [21] R. J. P. DeFigueiredo, W. R. Shankle, A. Maccato et al., “Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 12, pp. 5530–5534, 1995.
- [22] J. Ramírez, J. M. Górriz, F. Segovia et al., “Computer aided diagnosis system for the Alzheimer’s disease based on partial least squares and random forest SPECT image classification,” *Neuroscience Letters*, vol. 472, no. 2, pp. 99–103, 2010.
- [23] R. Chaves, J. Ramírez, J. M. Górriz et al., “SVM-based computer-aided diagnosis of the Alzheimer’s disease using t-test NMSE feature selection with feature correlation weighting,” *Neuroscience Letters*, vol. 461, no. 3, pp. 293–297, 2009.
- [24] P. Padilla, J. M. Górriz, J. Ramírez et al., “Analysis of SPECT brain images for the diagnosis of Alzheimer’s disease based on NMF for feature extraction,” *Neuroscience Letters*, vol. 479, no. 3, pp. 192–196, 2010.
- [25] F. Segovia, J. M. Górriz, J. Ramírez et al., “Classification of functional brain images using a GMM-based multi-variate approach,” *Neuroscience Letters*, vol. 474, no. 1, pp. 58–62, 2010.
- [26] S. Duchesne, A. Caroli, C. Geroldi, C. Barillot, G. B. Frisoni, and D. L. Collins, “MRI-based automated computer classification of probable AD versus normal controls,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 509–520, 2008.
- [27] M. R. Daliri, “Automated diagnosis of alzheimer disease using the scale-invariant feature transforms in magnetic resonance images,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 995–1000, 2012.
- [28] M. Graña, M. Termenon, A. Savio et al., “Computer Aided Diagnosis system for Alzheimer Disease using brain Diffusion Tensor Imaging features selected by Pearson’s correlation,” *Neuroscience Letters*, vol. 502, no. 3, pp. 225–229, 2011.
- [29] Z. Pawlak, “Rough sets,” *International Journal of Computer & Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [30] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, vol. 9, Springer, 1991.
- [31] Z. Pawlak, “Rough set approach to knowledge-based decision support,” *European Journal of Operational Research*, vol. 99, no. 1, pp. 48–57, 1997.
- [32] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Menlo Park, Calif, USA, 1989.
- [33] F. Ahmad, N. A. Mat Isa, Z. Hussain, and M. K. Osman, “Intelligent medical disease diagnosis using improved hybrid genetic algorithm—multilayer perceptron network,” *Journal of Medical Systems*, vol. 37, no. 2, article 9934, 2013.
- [34] M. R. Daliri, “A hybrid automatic system for the diagnosis of lung cancer based on genetic algorithm and fuzzy extreme learning machines,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 1001–1005, 2012.
- [35] J. Peral, *Probabilistic Reasoning in Intelligent Systems*, vol. 12, Morgan Kaufmann, San Mateo, Calif, USA, 1988.
- [36] C. B. Hall, J. Ying, L. Kuo, and R. B. Lipton, “Bayesian and profile likelihood change point methods for modeling cognitive function over time,” *Computational Statistics & Data Analysis*, vol. 42, no. 1, pp. 91–109, 2003.
- [37] P. Bélisle, L. Joseph, D. B. Wolfson, and X. Zhou, “Bayesian estimation of cognitive decline in patients with Alzheimer’s disease,” *The Canadian Journal of Statistics*, vol. 30, no. 1, pp. 37–54, 2002.
- [38] M. J. Prince, “Predicting the onset of Alzheimer’s disease using Bayes’ theorem,” *American Journal of Epidemiology*, vol. 143, no. 3, pp. 301–308, 1996.
- [39] R. W. Swiniarski and A. Skowron, “Rough set methods in feature selection and recognition,” *Pattern Recognition Letters*, vol. 24, no. 6, pp. 833–849, 2003.
- [40] A. Skowron and C. Rauszer, “The discernibility matrices and functions in information systems,” in *Intelligent Decision Support*, pp. 331–362, Springer, 1992.
- [41] C.-S. Ong, J.-J. Huang, and G.-H. Tzeng, “Building credit scoring models using genetic programming,” *Expert Systems with Applications*, vol. 29, no. 1, pp. 41–47, 2005.
- [42] W.-Y. Liang and C.-C. Huang, “The generic genetic algorithm incorporates with rough set theory—an application of the web services composition,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 5549–5556, 2009.
- [43] T. E. McKee and T. Lensberg, “Genetic programming and rough sets: a hybrid approach to bankruptcy classification,” *European Journal of Operational Research*, vol. 138, no. 2, pp. 436–451, 2002.
- [44] H. I. Elshazly, N. I. Ghali, A. M. E. Korany, and A. E. Hassanien, “Rough sets and genetic algorithms: a hybrid approach to breast cancer classification,” in *Proceedings of the World Congress on Information and Communication Technologies (WICT’12)*, pp. 260–265, IEEE, Trivandrum, India, November 2012.
- [45] H. Shimodaira, “A new genetic algorithm using large mutation rates and population-elitist selection (GALME),” in *Proceedings Eighth IEEE International Conference on Tools with Artificial Intelligence (ICTAI ’96)*, 1996.
- [46] S. Gauthier, B. Reisberg, M. Zaudig et al., “Mild cognitive impairment,” *The Lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.
- [47] D. W. Molloy and T. I. M. Standish, “A guide to the standardized Mini-Mental State Examination,” *International Psychogeriatrics*, vol. 9, no. 1, pp. 87–94, 1997.
- [48] A. Aggarwal and E. Kean, “Comparison of the folstein mini mental state examination (MMSE) to the montreal cognitive assessment (MoCA) as a cognitive screening tool in an inpatient rehabilitation setting gradient plasticity,” *Neuroscience & Medicine*, vol. 1, no. 2, 2010.

- [49] D. Mungas, "In-office mental status testing: a practical guide," *Geriatrics*, vol. 46, no. 7, pp. 54–66, 1991.
- [50] T. Smith, N. Gildeh, and C. Holmes, "The Montreal cognitive assessment: validity and utility in a memory clinic setting," *Canadian Journal of Psychiatry*, vol. 52, no. 5, pp. 329–332, 2007.
- [51] I. de Koning, F. van Kooten, and P. J. Koudstaal, "Value of screening instruments in the diagnosis of post-stroke dementia," *Pathophysiology of Haemostasis and Thrombosis*, vol. 28, no. 3-4, pp. 158–166, 1998.
- [52] P. Robert, S. Ferris, S. Gauthier, R. Ihl, B. Winblad, and F. Tennigkeit, "Review of Alzheimer's disease scales: is there a need for a new multi-domain scale for therapy evaluation in medical practice?" *Alzheimer's Research and Therapy*, vol. 2, no. 4, article 24, 2010.