Chapter 4

Experiment Results

In this chapter, we will show our experimental results, including anatomical differences between two different groups with voxel-based morphometric analysis and predictions on real data in each neurological disease. In VBM analysis, three statistical models are separately presented based on grey matter, white matter and cerebrospinal fluid. According to each statistical analysis, a corresponding classifier is trained with real data. Therefore, three classifiers are constructed and then combined to make a prediction on a subject in our system.

4.1 Materials

In our work, three study groups are gathered by Taipei Veterans General Hospital, including healthy people, patients with spinocerebellar ataxia type 3 (SCA3) and patients carrying bipolar disorder (BD). Demographic and clinical data of all subjects are summarized in Table 4.1, where the international cooperative ataxia rating scale (ICARS) [31] is a pharmacological assessment of the cerebellar syndrome and is performed on SCA3 patients. In addition, all healthy controls were diagnosed without carrying SCA3 and BD by doctors. Also, they are not members of SCA3 and BD patients' families.

Table 4.1: Demographic and clinical data of three study groups.

characteristics	Healthy controls			SCA3 patients			BD patients		
	$M+F$	M	F	$M+F$	M	F	$M+F$	M	F
amount(n)	76	34	42	6	3	3	15	5	10
age years(mean)	28.04	26.58	29.22	44.33	40.33	48.33	37.2	39.6	36
ICARS score	-	-	-	22	19	25	-	-	-

Magnetic resonance images of SCA3 patients were acquired from the 1.5T Siemens scanner at the Taipei Veterans General Hospital with a T1-weighted 3D IR sequence with TR = 9.7ms, TE = 4ms, FA = 12°, matrix size = 512×512, slices = 160, voxel size =

$0.47 \times 0.47 \times 1mm^3$. On the other hand, magnetic resonance images of the others were acquired from the 1.5T GE scanner at the Taipei Veterans General Hospital with TR = 8.672ms, TE=1.86ms, FOV = $26 \times 26 \times 10cm^3$, matrix size = 256×256, slices = 124, voxel size = $1.02 \times 1.02 \times 1.5mm^3$. It is obvious that the image quality from the GE scanner is better than that from the Siemens. Thus, as mentioned above, two segmentation tools would be used according to the image quality. All of the volume data were originally saved in DICOM format and were transformed into ANALYZE format before our procedure. Then, VBM analysis and classifier training were applied on both two neurological diseases respectively.

4.2 Structural Analysis for Patients Suffering Spinocerebellar Ataxia Type 3

Spinocerebellar ataxia (SCA) is a genetic disease and is classified into more than 26 types. Spinocerebellar ataxia type 3 (SCA3) is one of that and is an autosomal dominant neuromuscular degenerative disorder, also known as Machado-Joseph disease, named for affected families of Azorean extraction. Its clinical characteristics are progressive gait, limb ataxia, dysarthria, pyramidal signs, oculomotor disorders and degeneration of the cerebellum, the spinal cord and the brain stem [32]. Sadly, there is still no remedy for eradicating SCA but for softening symptoms. Here, an experiment was designed to analyze the brain structural differences between patients with SCA3 and normal subjects by VBM method and then to build up a diagnosis system for SCA3.

A study group was composed of six patients carrying SCA3 and eighteen normal subjects scanned by 1.5T Siemens scanner at the Taipei Veterans General Hospital. MR images of all subjects were normalized, segmented and modulated by SPM2 software. Here, we did not apply FSL software to segment all MR images because segmentation effects by SPM2

software were better than by FSL software on those images generated by 1.5T Siemens scanner with low resolution. Hence, eighteen normal subjects were not included in following classification process but six patients were included in the following classification stage. After being smoothed with 8mm FWHM isotropic Gaussian kernel, all preprocessed images were analyzed brain discrepancies by a VBM method with a two sample *t*-test. For the final *t*-map, the significance level was set at $p < 0.00005$ uncorrected for grey matter volume atrophy and at $p < 0.001$ uncorrected for white matter volume atrophy and CSF volume increase. An extension threshold was set 40 voxels to limit the minimum cluster size when showing results.

Figure 4.1 and Figure 4.2 illustrate locations of grey matter volume loss and white matter volume loss in SCA3 patients compared with normal controls by VBM method. Anatomical interpretations of these detected significant locations are summarized in Table 4.2 and Table 4.3 corresponding to Figure 4.1 and Figure 4.2 individually. These anatomical interpretations were processed by the software Talairach Daemon Client which inputs a brain coordinate obtained by VBM analysis and outputs a consistent anatomical label. Since CSF situates at the rim of the whole brain, the validity of applying VBM manner on CSF is influenced by the segmentation technique and is more incorrect than that of GM or WM. So, a VBM analysis of CSF was still experimented, shown in Figure 4.3, and was for verifying the VBM analysis of GM and WM. There are some inconsistent between VBM results of CSF and those of GM or WM. Hence, features extracted from CSF were not used in classification.

Applying VBM method on comparing SCA3 patients to normal subjects, the results revealed symmetrically significant atrophy of the anterior and posterior lobes of cerebellar hemispheres, basal ganglia (including lentiform nucleus, caudate nucleus, putamen), frontal lobe (including inferior, middle, superior and precentral frontal gyrus) and brainstem. Neurodegeneration of the cerebellum and brainstem stands to reason. Basal ganglis is associated with motor and learning functions. Frontal areas are concerned about cog-

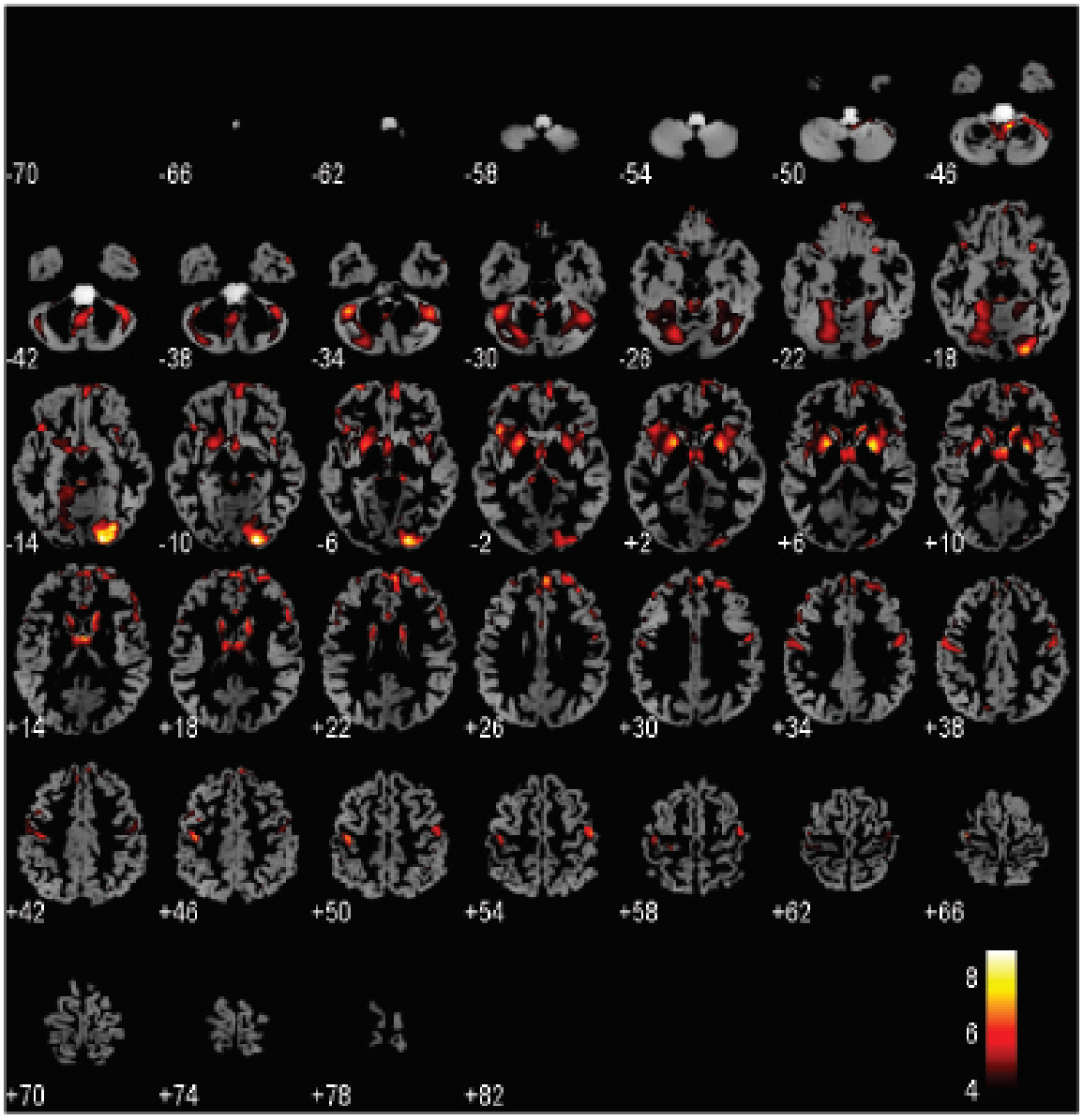


Figure 4.1: **Volumetric atrophy of grey matter in SCA3 patients by VBM analysis method.** Those detected regions where represent significant differences of volume decrease between SCA3 patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa. More detailed information of these areas are listed in Table 4.2.

Table 4.2: **Anatomical interpretation of grey matter volumetric atrophy in SCA3 patients.** These detected atrophy regions are separately presented according as cerebrum, cerebellum and brain stem and are sorted by their peak *t* values of each cluster. The *t* value reveals the significance of brain atrophy at the location in SCA3 patients. A cluster may be so large that there is more than one peak inside it.

Location	Side	Talairach coordinate(mm)			Peak t value	Cluster size
		x	y	z		
Cerebrum						
Occipital Lobe, Lingual Gyrus	R	21	-90	-4	10.42	7179
Occipital Lobe, Inferior Occipital Gyrus (BA18)	R	29	-86	-9	8.44	
Sub-lobar, Lentiform Nucleus, Putamen	R	25	3	7	8.21	4998
Frontal Lobe, Inferior Frontal Gyrus (BA47)	R	43	14	-4	5.98	
Sub-lobar, Extra-Nuclear (BA13)	R	40	8	-9	5.86	
Frontal Lobe, Inferior Frontal Gyrus	L	-42	17	-3	7.85	2295
Sub-lobar, Insula	L	-31	20	0	5.15	
Frontal Lobe, Superior Frontal Gyrus (BA9)	R	6	55	22	7.83	7930
Inter-Hemispheric		0	58	16	6.81	
Frontal Lobe, Middle Frontal Gyrus	R	37	44	21	6.4	
Sub-lobar, Lentiform Nucleus, Putamen	L	-20	7	4	7.71	5259
Sub-lobar, Lateral Ventricle	R	3	-2	12	7.51	4382
Limbic Lobe, Anterior Cingulate	R	1	2	-9	6.34	
Frontal Lobe, Precentral Gyrus (BA4)	L	-41	-16	44	7.43	2450
Frontal Lobe, Superior Frontal Gyrus (BA10)	L	-29	62	-7	7.43	303
Frontal Lobe, Precentral Gyrus	L	-36	-13	61	7.28	132
Sub-lobar, Caudate	R	12	21	7	7.18	2086
Frontal Lobe, Precentral Gyrus	R	50	-11	49	6.75	952
Frontal Lobe, Middle Frontal Gyrus	L	-42	39	28	6.32	159
Frontal Lobe, Precentral Gyrus (BA6)	R	50	-9	33	6.27	1007
Frontal Lobe, Inferior Frontal Gyrus	R	33	18	-15	6.19	406
Frontal Lobe, Middle Frontal Gyrus (BA46)	R	51	40	14	5.42	
Frontal Lobe, Superior Frontal Gyrus (BA11)	R	23	41	-22	5.92	435
Sub-lobar, Caudate	L	-11	12	16	5.92	1108
Inter-Hemispheric, Corpus Callosum		0	22	16	5.66	344
Cerebellum						
Posterior Lobe, Cerebellar Tonsil	R	9	-47	-35	8.18	10576
Posterior Lobe, Declive	L	-21	-70	-20	5.92	
Anterior Lobe, Culmen	L	-41	-47	-25	7.62	1785
Anterior Lobe, Culmen	R	43	-49	-26	6.87	3755
Brainstem						
Midbrain	L	-2	-34	-9	6.27	
Midbrain	R	16	-26	-5	6.14	283

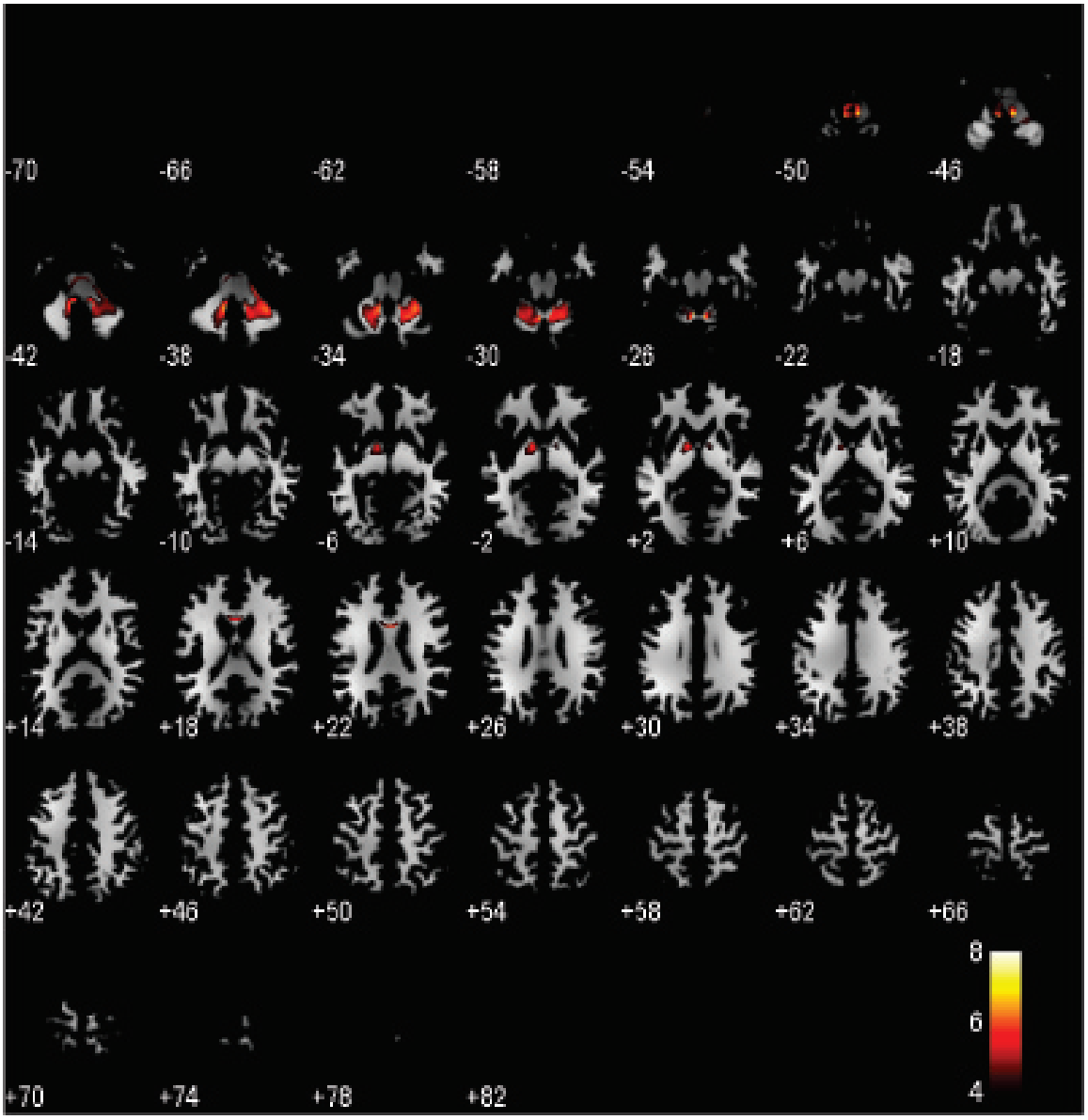


Figure 4.2: **Volumetric atrophy of white matter in SCA3 patients by VBM analysis method.** Those detected regions where represent significant differences of volume atrophy between SCA3 patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa. More detailed information of these areas are listed in Table 4.3.

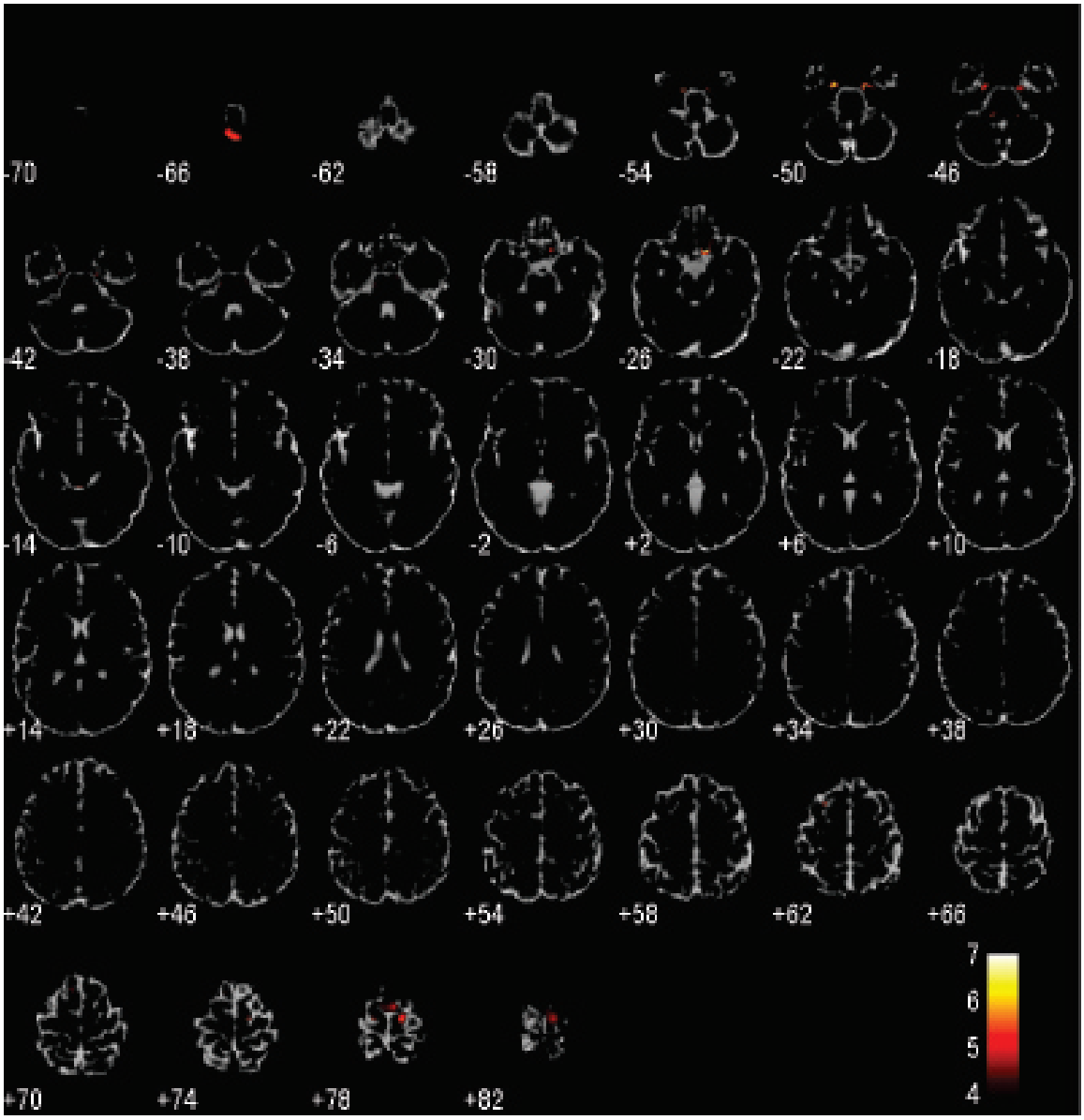


Figure 4.3: **Volumetric enlargement of CSF in SCA3 patients by VBM analysis method.** Those detected regions where represent significant differences of volume increase between BD patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa.

Table 4.3: **Volumetric atrophy of white matter in SCA3 patients by VBM analysis method.** These detected atrophy regions are separately presented according as cerebrum, cerebellum and brainstem and are sorted by their peak t values of each cluster. The t value reveals the significance of brain atrophy at the location in SCA3 patients. A cluster may be so large that there is more than one peak inside it.

Location	Side	Talairach coordinate(mm)			Peak t value	Cluster size
		x	y	z		
Cerebrum						
Sub-lobar, Lentiform Nucleus	L	-11	2	-1	5.93	2856
Inter-Hemispheric, Corpus Callosum		1	12	18	5.56	1207
Sub-lobar Extra-Nuclear, Corpus Callosum	R	4	5	21	4.98	
Frontal Lobe, Inferior Frontal Gyrus	R	25	18	-12	4.66	227
Sub-lobar, Lentiform Nucleus	R	11	2	2	4.56	1380
Cerebellum						
Anterior Lobe, Fastigium	L	-4	-54	-20	8.39	20162
Anterior Lobe, Fastigium	R	6	-55	-20	7.4	
Brainstem						
Medulla	R	8	-39	-38	7.81	

nitive functions and may be affected by cerebellar dysfunction. Also, some literature has confirmed that these areas are pathologically relative to SCA3 [32, 33]. Moreover, there were two locations detected in our results: the right occipital lobe and the lateral ventricle. There were some clinical vision problems in our patients which may lead to a significant atrophy in the occipital lobe. The unreasonable finding in the lateral ventricle may be caused by an incorrect registration around the thalamus.

In the foloowing classification processing, MR images of 76 normal subjects scanned from the GE scanner and six SCA3 patients from the Siemens scanner were separately segmented into GM, WM and CSF images with FSL software and SPM2 software. Then,

GM/WM images were individually normalized to the customized GM/WM templates constructed in VBM analysis of SCA3. Then, these processed images were modulated to restore volume changes. A GM mask was created to make a ROI selection on modulated GM images of each subject by selecting voxels whose t value in GM t-map is over 4. After thresholding, the remainder voxels extracted from GM images formed a new feature set for post-processing. A WM mask was also built up to make a ROI selection on modulated WM images of all subjects with a t-value threshold which exceeded 3.5. Equally, a new feature set was obtained by examining t values of all voxels in WM images. As a result of less validity of CSF analysis, we excluded the information from CSF for classification.

4.3 Structural Analysis for Patients Suffering Bipolar Disorder

Bipolar disorder (BD), also known as manic-depressive illness, is a kind of mood disorder that causes unusual shifts in a person's mood, energy, and ability to function. People with bipolar disorder experience mood episodes. These mood episodes can include depressive episodes, manic episodes, and mixed episodes. During depressive episodes, individuals usually experience sad mood, diminished interest in usual activities and disturbances in sleep, appetite, energy, and concentration. Manic episodes typically involve either extremely happy or irritable mood, accompanied by other changes in behavior, such as increased activity, decreased need for sleep, grandiose thinking, and racing thoughts. Mixed episodes involve the simultaneous occurrence of depressive and manic symptoms. The duration of mood episodes typically lasts from a couple of hours to many months. Between episodes people with BD often return to their usual functioning and personality. In addition, there are some related works inferring that BD has been associated with abnormalities of brain structure [15, 34]. Here, a experiment was designed to analyze the brain structural differences between patients with BD and normal subjects by VBM method and

then to build up a diagnosis system for bipolar disorder.

A study group was composed of fifteen patients carrying BD and thirty normal subjects who are selected from 76 normal subjects to be age-matched and gender-matched to the utmost. Demographic data of this study group are shown in Table 4.4. All MR images of these subjects were initially segmented with FSL software and then operated with serial preprocessing including normalization and modulation with SPM2 software. Before VBM analysis, all images were smoothed with an 8mm FWHM isotropic Gaussian kernel. Then, structural brain differences of preprocessed images were statistically analyzed between two groups by a VBM method with a two sample *t*-test. For each tissue, the significance level was set at $p < 0.001$ uncorrected for the final *t*-map and an extension threshold was set 40 voxels to limit the minimum cluster size when displaying results.

Table 4.4: Demographic and clinical data of BD study groups.

characteristics	Healthy controls			BD patients		
	$M+F$	M	F	$M+F$	M	F
amount(n)	30	14	16	15	5	10
age years(mean)	32.77	30.86	34.44	37.2	39.6	36

Figure 4.4, Figure 4.5 and Figure 4.6 illustrate combinative positions of grey matter volume loss, white matter volume growth and CSF volume growth in BD patients compared with normal subjects by VBM analysis respectively. Since some studies have shown the relationship between the ventricle size and BD patients, a VBM analysis of CSF was still experimented and was used to create a CSF classifier in the following process. Anatomical interpretations of those detected significant locations are summarized in Table 4.5 where reveals volume loss in BD patients and Table 4.6 where shows volume growth in BD patients. These explications were also obtained by the software Talairach Daemon Client.

Comparing brain structural differences between BD patients and normal subjects, significant volume loss was located at the frontal lobe (including precentral, inferior and mid-

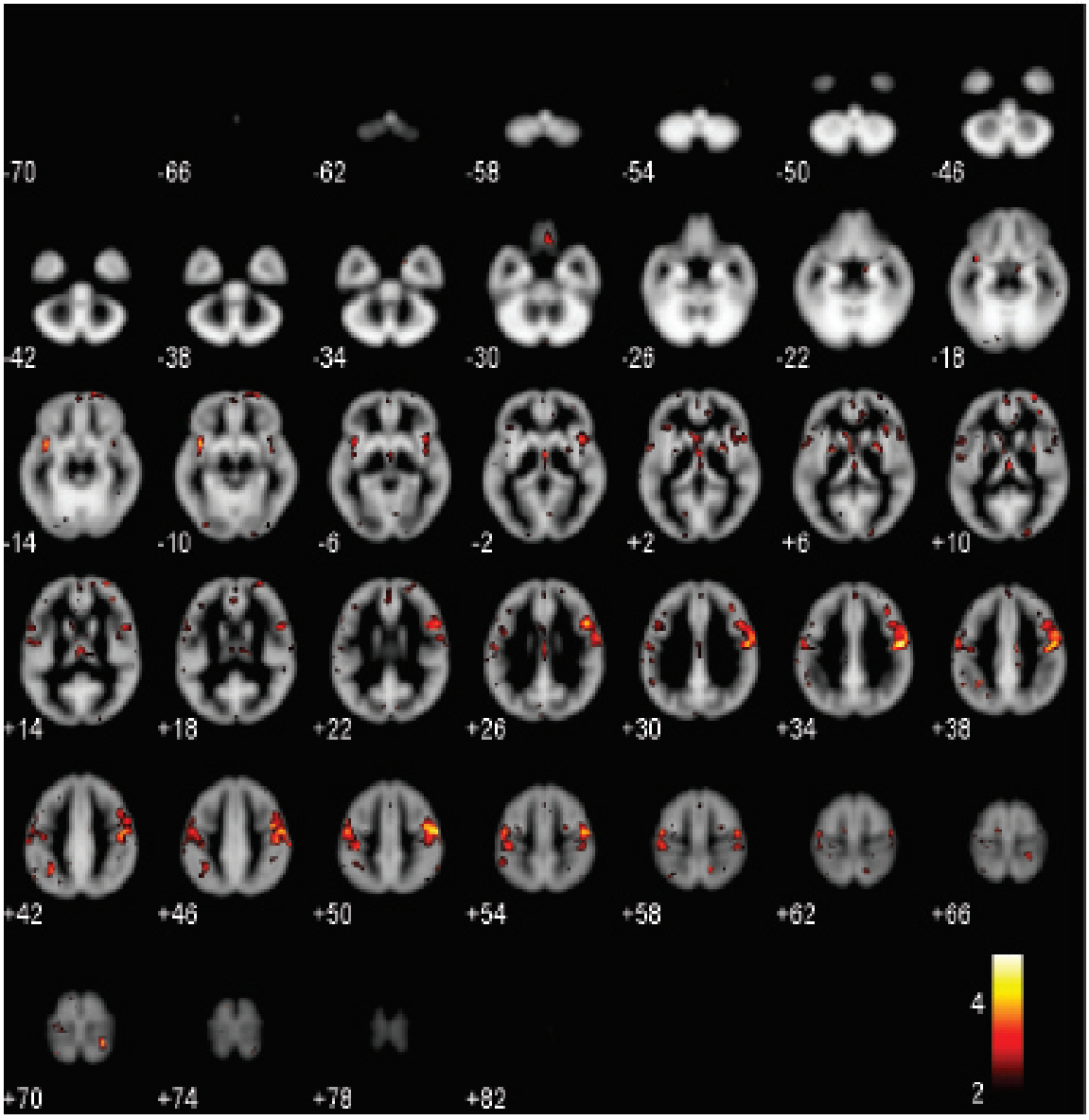


Figure 4.4: **Volumetric atrophy of gray matter in BD patients by VBM analysis method.** Those detected regions where represent significant differences of volume atrophy between BD patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa. More detailed information of these areas are listed in Table 4.5.

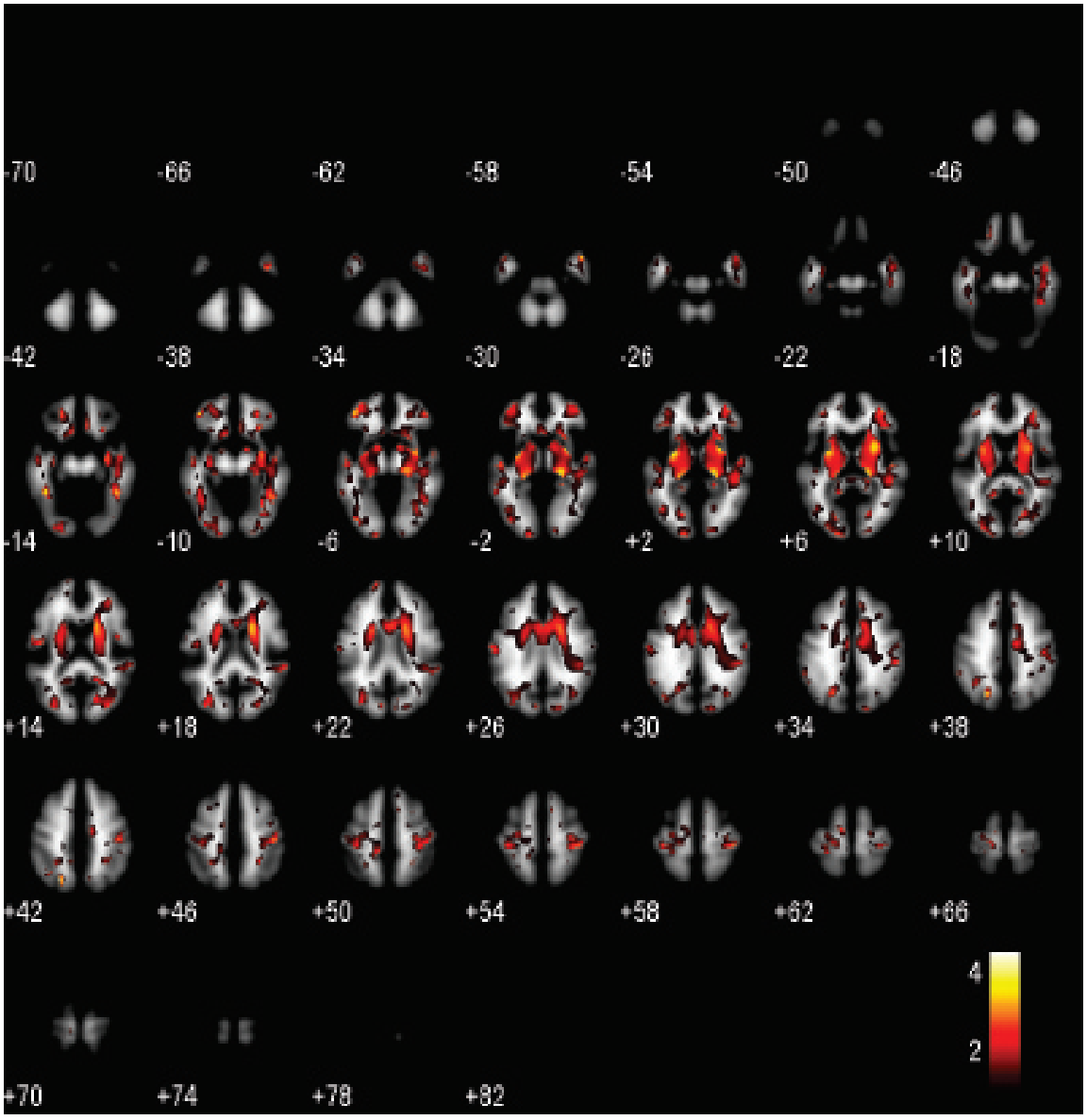


Figure 4.5: **Volumetric enlargement of white matter in BD patients by VBM analysis method.** Those detected regions where represent significant differences of volume expansion between BD patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa. More detailed information of these areas are listed in Table 4.6.

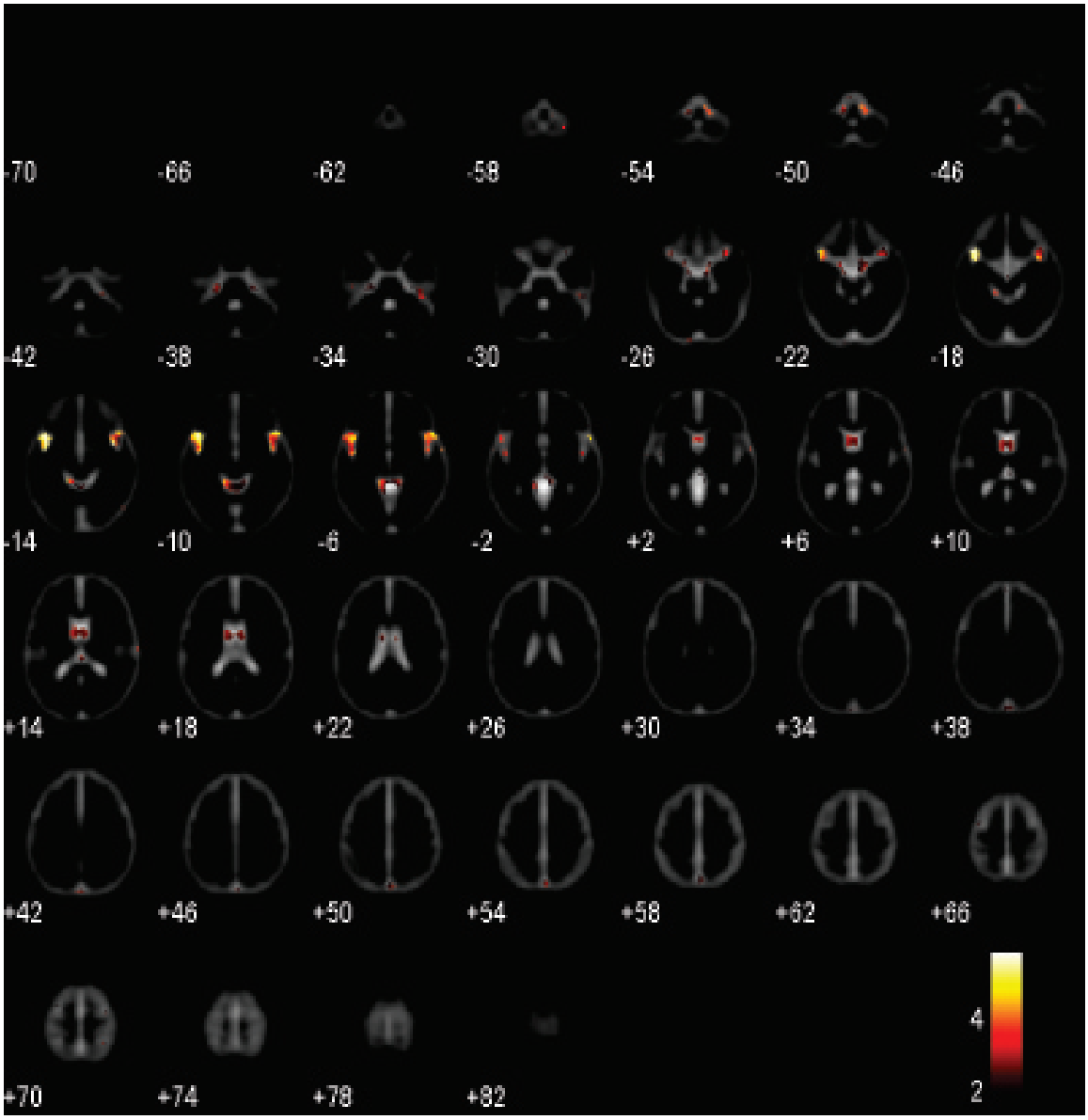


Figure 4.6: **Volumetric enlargement of CSF in BD patients by VBM analysis method.** Those detected regions where represent significant differences of volume increase between BD patients and normal subjects are colored from black to white (maximum) via reds and yellows according to their significances. Higher significance is displayed with brighter color, vice versa. More detailed information of these areas are listed in Table 4.6.

Table 4.5: **Brain structural atrophy in BD patients by VBM analysis method.** These detected atrophy regions are sorted by their peak *t* values of each cluster. The *t* value reveals the significance of brain atrophy at the location in BD patients. A cluster may be so large that there is more than one peak inside it.

Location	Side	Talairach coordinate(mm)			Peak t value	Cluster size
		x	y	z		
Cerebrum						
Frontal Lobe, Precentral Gyrus	R	54.45	-7.1557	31.6772	5.160	4567
Frontal Lobe, Precentral Gyrus (BA6)	R	42.57	-1.6681	44.2987	4.270	
Frontal Lobe, Inferior Frontal Gyrus	R	49.5	15.7273	23.1629	4.270	641
Sub-lobar, Extra-Nuclear (BA13)	L	-38.61	7.2465	-10.4552	4.160	305
Frontal Lobe, Precentral Gyrus (BA4)	L	-48.51	-7.4348	45.5085	4.060	532
Frontal Lobe, Precentral Gyrus	L	-41.58	-8.9586	53.8751	3.840	
Sub-lobar, Lateral Ventricle	L	-3.96	17.6221	2.8028	3.980	40
Parietal Lobe, Postcentral Gyrus	L	-42.57	-23.5823	52.7646	3.770	313
Frontal Lobe, Middle Frontal Gyrus (BA10)	R	30.69	61.6314	8.8908	3.750	81
Parietal Lobe, Inferior Parietal Lobule	L	-34.65	-52.4129	39.4689	3.730	102
Frontal Lobe, Precentral Gyrus	L	-58.41	-4.0655	35.2072	3.640	86
Sub-lobar, Insula (BA13)	R	42.57	12.5103	-2.3081	3.580	59
Sub-lobar, Third Ventricle	R	1.98	-2.9483	-0.6935	3.530	44
Frontal Lobe, Middle Frontal Gyrus	R	43.56	14.2955	33.3672	3.450	73

dle frontal gyrus), the parietal lobe (including postcentral and inferior parietal gyrus), the insula, the lateral ventricle and the third ventricle. Several researches show that phenomena of concentration or volume loss in frontal lobe are appeared in brain structures of BD patients [35, 36]. Moreover, the findings of postcentral gyrus in the parietal lobe are sensible because it is concerned with sensory fields. However, two unreasonable results, the lateral ventricle and the third ventricle, were presented and inconsistent with literature. The results may be caused by small sample population so that each subject has large contribution to the statistical analysis. In the other hand, significant volume increase was detected in Temporal

Table 4.6: **Brain structural enlargement in BD patients by VBM analysis method.** These detected increase regions are sorted by their peak t values of each cluster. The t value reveals the significance of brain enlargement at the location in BD patients. A cluster may be so large that there is more than one peak inside it.

Location	Side	Talairach coordinate(mm)			Peak t value	Cluster size
		x	y	z		
Cerebrum						
Temporal Lobe, Inferior Frontal Gyrus	R	36.63	8.1313	-12.18	5.870	2184
Frontal Lobe, Inferior Frontal Gyrus	R	45.54	20.8515	-10.30	5.550	
Frontal-Temporal Space	R	54.45	11.5835	-1.42	4.940	
Temporal Lobe, Superior Temporal Gyrus (BA38)	L	-37.62	8.0054	-14.70	5.850	3604
Frontal Lobe, Inferior Frontal Gyrus	L	-38.61	15.7557	-15.09	5.850	
Temporal Lobe, Superior Temporal Gyrus	L	-40.59	-0.5458	-10.91	5.540	
Temporal Lobe, Middle Temporal Gyrus (BA21)	R	62.37	-0.3359	-6.71	4.630	14
Parietal Lobe, Precuneus	L	-23.76	-65.9299	41.0664	4.550	157
Temporal Lobe, Superior Temporal Gyrus	L	-58.41	1.9376	-0.097	4.55	22
Temporal Lobe, Superior Temporal Gyrus	R	31.68	17.1895	-25.2507	4.43	51
Temporal Lobe, Fusiform Gyrus	L	-40.59	-47.1735	-11.0962	4.250	63
Sub-lobar, Thalamus, Pulvinar	R	19.8	-28.9257	4.2109	4.250	342
Temporal Lobe, Superior Temporal Gyrus	R	63.36	-6.4137	7.6902	3.95	3
Frontal Lobe, Middle Frontal Gyrus	L	-39.6	40.3532	-8.7478	3.720	57
Sub-lobar, Lentiform Nucleus, Putamen	L	-25.74	-2.8144	1.9831	3.710	208
Sub-lobar, Lentiform Nucleus, Putamen	R	26.73	3.1363	4.4488	3.530	180
Sub-lobar, Extra-Nuclear	R	27.72	0.8009	-3.4043	3.440	
Sub-lobar, Extra-Nuclear	R	19.8	8.532	15.2326	3.380	48
Cerebellum						
Anterior Lobe, Culmen	L	-13.86	-35.2962	-6.64	5.470	574

lobe (including inferior, superior and middle temporal gyrus), the frontal lobe (including inferior and middle frontal gyrus), the thalamus, basal ganglia (including lentiform nucleus and putamen) and the left cerebellum anterior lobe. Some works revealed the volume increase of temporal lobe, thalamus and basal ganglia though some have opposite comments [37, 38]. Nevertheless, there is no clear pathological discovery in brain structures of BD patients. Therefore, these detected significant regions were entirely used for ROI selection and extracted features were all for classification.

In this work, MR images of 76 normal subjects and fifteen BD patients were partitioned into GM, WM and CSF images, normalized to the customized templates (including GM template, WM template and CSF template) constructed in VBM analysis of bipolar disorder and modulated to correct volume changes respectively. GM, WM and CSF masks were separately created to make ROI selections on modulated GM, WM and CSF images by selecting voxels whose t value in individual t-maps are all over 1.68. After thresholding, three new feature sets were formed apart according to GM, WM and CSF images and were used for classification.

4.4 Experiments on Diagnosis System

For a particular disease, there are three feature sets according to GM, WM and CSF models. For each model, the feature set was applied PCA to find proper representations for it with fewer variables. Then, both of variance-based principal component selection and significant-based principal component selection were used on the reduced feature set to choose more useful features to build up a suitable classification space. In our experiments, we used a leave-one-out cross validation to verify our work. That is, the whole subjects were separated into two parts: a test subject and training subjects. Each of all subjects was picked out to be the test subject and the others composed training set to train a proper classifier and to make a prediction on the test subject.

Table 4.7: **Predictions by a SCA3 classifier on SCA3 patients with variance-based PC selection method.** These test subjects are composed of 6 SCA3 patients diagnosed by doctors and 76 normal subjects. The columns of GM prediction and WM prediction represent the possibilities for test subjects of falling ill which are estimated by the GM classifier and by the WM classifier respectively. The column of result represents the final prediction on test subjects and is obtained by choosing the maximum probability of being abnormal from the results of the GM and WM classifiers.

SCA3 Patients	GM prediction (%)	WM prediction (%)	Result (%)
1	0.999995	0.695823	0.999995
2	1	0.505933	1
3	1	0.916646	1
4	1	0.737166	1
5	0.653248	0.782884	0.782884
6	0.99247	0.784765	0.99247

4.4.1 Results of SCA3 Diagnosis System

In SCA3 classification model, 76 normal subjects and six SCA3 patients were used to train the SCA3 classifier and to validate the performance of the SCA3 classifier. As mentioned above, we only constructed a GM classifier and a WM classifier and then combined them together without constructing a CSF classifier because of its less validity. In this system, the prior parameter was set as 0.073 and the window size parameter was set as 4 for both of GM and WM classifiers.

With a variance-based PC selection method, a better performance was occurred when the variance ratio was 30% about 4 to 5 principal components in the GM classifier and was 40% about 2 PCs in the WM classifier. Table 4.7, Table 4.8 and Table 4.9 show the prediction results of all these subjects by using a variance-based PC selection method. The FN and FP rates are 0 and 0.013157895 respectively. That is, there is a false alarm in 76 normal subjects and the others are classified correctly.

Table 4.8: **Predictions by a SCA3 classifier on normal subjects with variance-based PC selection method.**

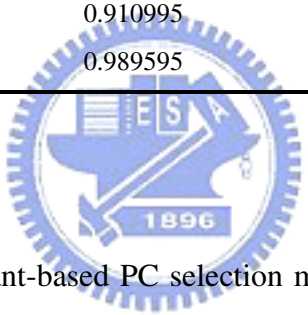
Controls	GM prediction (%)	WM prediction (%)	Result (%)
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0.000546	0.000546
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0.000567	0.000567
14	0	0	0
15	0	0.095555	0.095555
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	0	0	0
24	0	0	0
25	0	0	0
26	0	0	0
27	0	0	0
28	0	0	0
29	0	0	0
30	0	0	0
31	0	0	0
32	0	0	0
33	0	0	0
34	0	0	0
35	0	0	0
36	0	0	0
37	0	0	0
38	0	0	0

Table 4.9: **Predictions by a SCA3 classifier on normal subjects with variance-based PC selection method.** The colored row represents a wrong prediction.

Controls	GM prediction (%)	WM prediction (%)	Result (%)
39	0	0	0
40	0	0	0
41	0	0	0
42	0	0.000016	0.000016
43	0	0.000143	0.000143
44	0	0	0
45	0	0.148336	0.148336
46	0	0	0
47	0	0	0
48	0	0	0
49	0	0	0
50	0	0.083441	0.083441
51	0	0	0
52	0	0	0
53	0	0	0
54	0	0.666251	0.666251
55	0	0	0
56	0	0	0
57	0	0	0
58	0	0	0
59	0	0	0
60	0	0	0
61	0	0	0
62	0	0	0
63	0	0	0
64	0	0	0
65	0	0	0
66	0	0	0
67	0	0	0
68	0	0	0
69	0	0	0
70	0	0	0
71	0	0	0
72	0	0	0
73	0	0	0
74	0	0	0
75	0	0	0
76	0	0	0

Table 4.10: **Predictions by a SCA3 classifier on SCA3 patients with significant-based PC selection method.** These test subjects are composed of 6 SCA3 patients diagnosed by doctors and 76 normal subjects. The columns of GM prediction and WM prediction represent the possibilities for test subjects of falling ill which are estimated by the GM classifier and by the WM classifier respectively. The column of result represents the final prediction on test subjects and is obtained by choosing the maximum probability of being abnormal from the results of the GM and WM classifiers.

SCA3 Patients	GM prediction (%)	WM prediction (%)	Result (%)
1	0.999999	1	1
2	1	0.177831	1
3	1	0.999975	1
4	1	1	1
5	0.910995	0.999932	0.999932
6	0.989595	1	1



As adopting a significant-based PC selection method, a better performance was happened when the variance ratio was 20% about 2 to 3 principal components in the GM classifier and was 30% about 4 PCs in the WM classifier. Table 4.10, Table 4.11 and Table 4.12 display the prediction results of all these subjects by using a significant-based PC selection method. Both of the FN and FP rates are 0. That is, all subjects are classified correctly.

Comparing the effect of using variance-based PC selection method and that of using significant-based PC selection method, it reveals that the performance of the latter, 100% classification accuracy, is better than that of the former, 98.7% classification accuracy. It shows that methods of PC selection affect the efficiency of a classifier and the method of significant-based selection establishes a good space to represent data of two different groups and to distinguish them. Moreover, using significant-based PC selection method provides a more accurate possibility to predict whether a test subject is abnormal or not.

Table 4.11: **Predictions by a SCA3 classifier on normal subjects with significant-based PC selection method.**

Controls	GM prediction (%)	WM prediction (%)	Result (%)
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0.000002	0.000002
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	0	0	0
24	0	0	0
25	0	0	0
26	0	0	0
27	0	0	0
28	0	0	0
29	0	0	0
30	0	0	0
31	0	0	0
32	0	0	0
33	0	0	0
34	0	0	0
35	0	0	0
36	0	0	0
37	0	0	0
38	0	0	0

Table 4.12: **Predictions by a SCA3 classifier on normal subjects with significant-based PC selection method.**

Controls	GM prediction (%)	WM prediction (%)	Result (%)
39	0	0	0
40	0	0	0
41	0	0	0
42	0	0.000001	0.000001
43	0	0	0
44	0	0	0
45	0	0.000001	0.000001
46	0	0	0
47	0	0	0
48	0	0	0
49	0	0	0
50	0	0.000055	0.000055
51	0	0	0
52	0	0	0
53	0	0	0
54	0	0	0
55	0	0	0
56	0	0	0
57	0	0	0
58	0	0	0
59	0	0	0
60	0	0	0
61	0	0	0
62	0	0	0
63	0	0	0
64	0	0	0
65	0	0.000004	0.000004
66	0	0	0
67	0	0	0
68	0	0	0
69	0	0	0
70	0	0.000001	0.000001
71	0	0	0
72	0	0	0
73	0	0	0
74	0	0	0
75	0	0	0
76	0	0	0

4.4.2 Results of BD Diagnosis System

In BD classification model, 76 normal subjects and fifteen BD patients were used to train the BD classifier and to validate the performance of the BD classifier. We constructed GM, WM and CSF classifiers separately and then combined them together to decide the final prediction on a test subject. In this system, the prior parameter was set as 0.165 and the window size parameter was set as 5 for all of GM, WM and CSF classifiers.

With a variance-based PC selection method, a better performance was occurred when the variance ratio was 60% about 36 principal components in the GM classifier, was 80% about 46 to 47 PCs in the WM classifier and was 50% about 12 to 13 PCs in the CSF classifier. Table 4.13, Table 4.14 and Table 4.15 show the prediction results of all these subjects by using a variance-based PC selection method. The FN and FP rates are 0 and 0.276315789 respectively. That is, there is 21 false alarms in 76 normal subjects and the others are classified correctly.

As adopting a significant-based PC selection method, a better performance was happened when the variance ratio was 10% about 2 to 3 principal components in the GM classifier, was 80% about 61 to 69 PCs in the WM classifier and was 60% about 40 to 45 PCs in the CSF classifier. Table 4.16, Table 4.17 and Table 4.18 display the prediction results of all these subjects by using a significant-based PC selection method. The FN and FP rates are 0 and 0.157894737 respectively. That is, there are 12 false alarms in 76 normal subjects and the others are classified correctly.

Comparing the efficiency with a variance-based PC selection method and that with a significant-based PC selection method, it reveals that the performance of the latter, 86.8% classification accuracy, is better than that of the former, 76.9% classification accuracy. Again, it proves that the classification accuracy is influenced by methods of PC selection. In short, a significant-based principal component selection is a good choice to use when constructing a classification space in our method.

Table 4.13: **Predictions by a BD classifier on patients with variance-based PC selection method.** These test subjects are composed of 15 BD patients diagnosed by doctors and 76 normal subjects. The columns of GM prediction, WM prediction and CSF prediction represent the possibilities for test subjects of falling ill which are estimated by the GM classifier, the WM classifier and the CSF classifier respectively. The column of result represents the final prediction on test subjects and is obtained by choosing the maximum probability of being abnormal from the results of the GM, WM and CSF classifiers.

BD patients	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
1	0.997847	0.038082	0.665953	0.997847
2	0.000000	0.999999	0.000030	0.999999
3	0.000004	0.996431	0.565577	0.996431
4	0.999591	0.999994	0.995726	0.999994
5	0.999947	1.000000	1.000000	1.000000
6	0.999994	0.996339	0.999999	0.999999
7	0.744795	0.013229	0.298132	0.744795
8	0.983158	0.967153	0.086069	0.983158
9	0.988316	0.999999	0.952298	0.999999
10	0.615422	0.808894	0.628006	0.808894
11	0.308268	0.004680	0.712419	0.712419
12	0.929654	0.987404	0.103738	0.987404
13	0.000000	0.999999	0.001344	0.999999
14	0.999500	1.000000	1.000000	1.000000
15	0.975914	0.999996	0.998801	0.999996

Table 4.14: **Predictions by a BD classifier on normal subjects with variance-based PC selection method.** The colored row represents a wrong prediction.

Controls	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
1	0.000021	0.024830	0.000000	0.024830
2	0.000000	0.000000	0.000311	0.000311
3	0.997544	0.000048	0.050370	0.997544
4	0.000000	0.000000	0.000020	0.000020
5	0.000000	0.995277	0.000003	0.995277
6	0.000000	0.000002	0.090636	0.090636
7	0.000016	0.000000	0.000005	0.000016
8	0.000001	0.099958	0.001160	0.099958
9	0.000000	0.000007	0.000757	0.000757
10	0.000000	0.953099	0.530111	0.953099
11	0.000000	0.012948	0.004020	0.012948
12	0.000000	0.253951	0.000148	0.253951
13	0.000000	0.008661	0.276420	0.276420
14	0.000000	0.096330	0.000000	0.096330
15	0.000000	0.254212	0.004773	0.254212
16	0.000020	0.003003	0.022718	0.022718
17	0.000000	0.000028	0.000000	0.000028
18	0.048727	0.005161	0.003119	0.048727
19	0.000000	0.000866	0.015513	0.015513
20	0.000000	0.981618	0.004701	0.981618
21	0.000000	0.000002	0.000000	0.000002
22	0.000000	0.000144	0.671633	0.671633
23	0.996421	0.876781	0.000000	0.996421
24	0.000013	0.002096	0.231537	0.231537
25	0.019012	0.000175	0.123455	0.123455
26	0.521506	0.866327	0.000475	0.866327
27	0.000002	0.003812	0.000046	0.003812
28	0.000000	0.643584	0.406514	0.643584
29	0.000000	0.938766	0.025012	0.938766
30	0.000000	0.069816	0.006015	0.069816
31	0.000000	0.000000	0.020852	0.020852
32	0.000000	0.000086	0.003349	0.003349
33	0.000000	0.000163	0.006694	0.006694
34	0.000282	0.999104	0.072890	0.999104
35	0.000005	0.011477	0.000138	0.011477
36	0.959596	0.069707	0.027910	0.959596
37	0.000001	0.031256	0.012908	0.031256
38	0.000056	0.971662	0.000000	0.971662

Table 4.15: **Predictions by a BD classifier on normal subjects with variance-based PC selection method.** The colored row represents a wrong prediction.

Controls	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
39	0.000000	0.023203	0.006153	0.023203
40	0.000000	0.000001	0.072337	0.072337
41	0.180453	0.001491	0.000747	0.180453
42	0.000099	0.000009	0.362696	0.362696
43	0.977943	0.006881	0.000009	0.977943
44	0.000002	0.000023	0.000000	0.000023
45	0.000030	0.000000	0.481911	0.481911
46	0.000000	0.998972	0.556900	0.998972
47	0.000086	0.000016	0.000038	0.000086
48	0.000000	0.100556	0.035933	0.100556
49	0.000000	0.098304	0.010466	0.098304
50	0.000012	0.000081	0.184439	0.184439
51	0.000000	0.228474	0.015191	0.228474
52	0.971146	0.000170	0.225743	0.971146
53	0.000000	0.057463	0.000000	0.057463
54	0.000013	0.005559	0.061305	0.061305
55	0.000000	0.000011	0.179125	0.179125
56	0.000001	0.531536	0.078312	0.531536
57	0.000000	0.990739	0.000106	0.990739
58	0.001144	0.126917	0.003102	0.126917
59	0.000000	0.822923	0.076024	0.822923
60	0.026953	0.984142	0.024461	0.984142
61	0.000000	0.000029	0.103551	0.103551
62	0.000032	0.002383	0.005452	0.005452
63	0.000018	0.000046	0.061843	0.061843
64	0.000000	0.006143	0.003452	0.006143
65	0.000001	0.775748	0.001361	0.775748
66	0.000004	0.001544	0.020145	0.020145
67	0.004439	0.026738	0.023286	0.026738
68	0.000000	0.015553	0.193120	0.193120
69	0.000030	0.000233	0.000000	0.000233
70	0.000000	0.004315	0.352926	0.352926
71	0.000000	0.003973	0.089708	0.089708
72	0.874916	0.051553	0.119379	0.874916
73	0.000000	0.439994	0.031998	0.439994
74	0.000000	0.471110	0.008337	0.471110
75	0.000000	0.214042	0.000586	0.214042
76	0.001937	0.000014	0.015965	0.015965

Table 4.16: **Predictions by a BD classifier on BD patients with significant-based PC selection method.** These test subjects are composed of 15 BD patients diagnosed by doctors and 76 normal subjects. The columns of GM prediction, WM prediction and CSF prediction represent the possibilities for test subjects of falling ill which are estimated by the GM classifier, by the WM classifier and by the CSF classifier respectively. The column of result represents the final prediction on test subjects and is obtained by choosing the maximum probability of being abnormal from the results of the GM, WM and CSF classifiers.

BD Patients	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
1	0.951285	0.024180	0.934230	0.951285
2	0.009554	0.999999	0.000232	0.999999
3	0.038517	0.997940	0.539584	0.997940
4	0.767455	0.999852	0.946834	0.999852
5	0.838052	0.998756	1.000000	1.000000
6	0.785975	0.966927	0.996268	0.996268
7	0.579724	0.826220	0.525341	0.826220
8	0.614239	0.638330	0.036379	0.638330
9	0.438554	1.000000	0.998335	1.000000
10	0.447763	0.897645	0.978042	0.978042
11	0.636093	0.762916	0.951610	0.951610
12	0.584977	0.985007	0.030686	0.985007
13	0.012618	0.993619	0.000449	0.993619
14	0.946892	1.000000	1.000000	1.000000
15	0.941403	0.999919	0.998202	0.999919

Table 4.17: **Predictions by a BD classifier on normal subjects with significant-based PC selection method.** The colored row represents a wrong prediction.

Controls	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
1	0.072117	0.000420	0.000000	0.072117
2	0.096522	0.000000	0.000001	0.096522
3	0.825768	0.000301	0.000822	0.825768
4	0.008250	0.000000	0.000000	0.008250
5	0.001669	0.115922	0.000000	0.115922
6	0.000007	0.000002	0.005136	0.005136
7	0.581750	0.000000	0.000016	0.581750
8	0.085075	0.000298	0.000080	0.085075
9	0.000000	0.040672	0.000000	0.040672
10	0.059370	0.000475	0.805460	0.805460
11	0.000536	0.003190	0.001913	0.003190
12	0.000000	0.877471	0.000294	0.877471
13	0.018858	0.003047	0.163927	0.163927
14	0.009439	0.002579	0.000000	0.009439
15	0.036538	0.030513	0.000143	0.036538
16	0.027767	0.000013	0.000598	0.027767
17	0.049546	0.000004	0.000000	0.049546
18	0.116265	0.025119	0.001920	0.116265
19	0.054027	0.000010	0.001025	0.054027
20	0.000000	0.483980	0.000019	0.483980
21	0.000011	0.000002	0.000000	0.000011
22	0.034698	0.000018	0.001129	0.034698
23	0.682494	0.085684	0.000000	0.682494
24	0.102495	0.000094	0.004696	0.102495
25	0.002145	0.000009	0.015016	0.015016
26	0.051138	0.113759	0.000045	0.113759
27	0.024380	0.000086	0.000000	0.024380
28	0.000083	0.973751	0.368119	0.973751
29	0.000093	0.078249	0.000176	0.078249
30	0.001131	0.001526	0.000566	0.001526
31	0.002829	0.000000	0.000172	0.002829
32	0.062694	0.000006	0.000060	0.062694
33	0.000250	0.001561	0.000795	0.001561
34	0.097401	0.430682	0.008879	0.430682
35	0.042760	0.000031	0.000152	0.042760
36	0.278202	0.004467	0.000003	0.278202
37	0.078412	0.000064	0.004380	0.078412
38	0.077996	0.330170	0.000000	0.330170

Table 4.18: **Predictions by a BD classifier on normal subjects with significant-based PC selection method.** The colored row represents a wrong prediction.

Controls	GM prediction (%)	WM prediction (%)	CSF prediction (%)	Result (%)
39	0.069429	0.004617	0.000013	0.069429
40	0.021939	0.000000	0.188631	0.188631
41	0.020114	0.000014	0.000250	0.020114
42	0.001536	0.000000	0.000128	0.001536
43	0.247769	0.003364	0.001624	0.247769
44	0.000030	0.000005	0.000000	0.000030
45	0.062321	0.000000	0.000538	0.062321
46	0.000372	0.857919	0.124341	0.857919
47	0.062323	0.000000	0.000000	0.062323
48	0.045648	0.009273	0.000571	0.045648
49	0.009056	0.000141	0.000346	0.009056
50	0.004976	0.000010	0.000216	0.004976
51	0.001145	0.963876	0.000032	0.963876
52	0.386725	0.000007	0.168682	0.386725
53	0.126456	0.992163	0.000006	0.992163
54	0.009375	0.004837	0.001693	0.009375
55	0.100840	0.000011	0.000098	0.100840
56	0.104275	0.009482	0.002134	0.104275
57	0.000251	0.102376	0.000011	0.102376
58	0.010635	0.912783	0.001457	0.912783
59	0.000000	0.336177	0.000491	0.336177
60	0.214691	0.026343	0.000255	0.214691
61	0.000000	0.000050	0.000504	0.000504
62	0.001820	0.003585	0.013185	0.013185
63	0.094691	0.000004	0.001570	0.094691
64	0.000269	0.000044	0.001128	0.001128
65	0.000251	0.000000	0.000186	0.000251
66	0.084014	0.001904	0.000306	0.084014
67	0.034299	0.004575	0.305676	0.305676
68	0.000314	0.208391	0.082601	0.208391
69	0.097364	0.000140	0.000000	0.097364
70	0.000008	0.018353	0.021745	0.021745
71	0.121095	0.000510	0.001519	0.121095
72	0.595191	0.000049	0.001461	0.595191
73	0.000000	0.000029	0.013162	0.013162
74	0.000064	0.109301	0.000084	0.109301
75	0.006847	0.905395	0.000005	0.905395
76	0.222983	0.000012	0.000347	0.222983

Comparing the classification results of SCA3 classifiers and BD classifiers, it was clear that the performance of SCA3 classifiers were better than that of BD classifiers although the population of SCA3 study group was less than that of BD study group. We inferred that it resulted from the explicit pathology of SCA3 that led to apparent brain structural changes. On the other hand, due to the implicit brain volume changes of BD patients compared with normal controls, subtle discrepancy might influence the VBM analysis and the classification result. Thus, the efficiency of BD classifiers was not as good as that of SCA3 classifiers. In other words, our proposed procedure is suitable to predicting a disease which results in clear brain volume changes.