Detect the Disrupted Brain Connectivity in Type-II Diabetes Patients

Yue Wei, Weiyan Xie
Department of Computer Science
The Hong Kong University of Science and Technology
{yweiaj,wxieai}@connect.ust.hk

Abstract

This project explores the disrupted brain connectivity in type-II diabetes mellitus (T2DM) patients before the patients show measurable cognitive impairment. The research question is whether the diabetic brain can be distinguished from the normal brain by the brain connectivity features and whether the brain connectivity features can indicate the time course of T2DM. Our experimental results show that using the brain connectivity features, our obtained model can classify diabetes and healthy patients with 92.5% accuracy, and predict the duration of T2DM with the error ± 1.9 years for the T2DM patients. Our model can also point out which brain regions are affected most in the diabetic brain, which can help the earlier detection and earlier prevention of disrupted brain connectivity influenced by the T2DM.

1 Introduction

The type-II diabetes mellitus (T2DM) is a common metabolic disorder characterized by chronic high blood glucose levels [1]. Long-term hyperglycemia may damage the brain and impair the cognitive function of T2DM patients [2]. It has been revealed that T2DM patients are highly susceptible to several types of cognitive impairments disease, especially Alzheimer Disease [1]. In recent years, plenty of studies uncover brain atrophy distributed in different gray matter regions of T2DM patients, including right Rolandic operculum [2, 3], right temporal lobes [2, 4, 3], anterior cingulate [4], hippocampus [5] and medial frontal lobes [4]. In addition to regional gray matter loss, it is also revealed that the white matter integrity and the structural brain connectivity is disrupted in T2DM patients [3, 6]. Neurological studies show that disruption in white matter tract diffusion characteristics of many regions [3] (e.g. the connections in the temporal lobe and hippocampus, connections between frontal and temporal lobes, connections with thalamus) is associated with impaired glucose metabolism and cognitive dysfunction.

However, as most studies apply structure information to diagnose diabetic brain focus on the patient with measurable cognitive impairment [7], it is unclear when the structural brain changes in T2DM patients become distinguishable [8]. More specifically, it is worth exploring whether the diabetes brain without clinical cognitive loss can be identified from normal healthy brain based on structural images. Moreover, as most previous works either focus on exploring regional gray matter or white matter changes [2, 4, 5, 3, 6, 9], it is unknown that whether gray matter or white matter fiber connections are more valuable in distinguishing diabetic brain.

In this study, we utilize machine learning methods to explore new solutions in automatic diabetic brain diagnosis along with the prediction on duration of the disease. **Since we focus on the disrupted brain connectivity in the T2DM patients before they show the cognitive impairment, all patients in our study are cognitively normal.** Furthermore, we examine the features that contribute most to the classification and prediction models to identify a few most vulnerable brain regions and white matter fiber tract connections that characterize a typical diabetic brain.

Final Project Report for MATH 5470: Statistical Machine Learning

	HC(n=47)	T2DM(n=47)	Statistics	P-value
Age(years)	55.6±8.4	57.4±8.0	t=-1.064	0.305
Sex(%)	25(47)	18(47)	χ^2=2.100	0.147
Formal education(years)	10.0±1.4	10.7±1.8	t=-5.447	0.022
BMI(kg/m^2)	20.8±1.4	24.5±2.0	t=-104.244	<0.001
T2DM duration(years)	NA	6.6±3.4	NA	NA
FBG(mmol/L)	NA	6.7±2.1	NA	NA
HbA1c(%)	NA	8.0±1.5	NA	NA

Table 1: Demographic data and clinical biochemical indicators of all subjects (FBG: Fasting Blood Glucose; NA: Not Applicable).

2 Data Acquisition and Preprocessing

2.1 Participants

In total 94 participants are recruited with complete written informed consent. 47 of the participants are T2DM patients recruited from the first affiliated hospital of Guangzhou university of Chinese medicine, while other 47 participants are healthy people (Healthy control: HC group). Diagnosis is conducted based on the standard criteria according to American Diabetes Association. Cognitive impairment is assessed by clinical doctors and patient with measurable cognitive impairment are **excluded**. Other exclusion criterions included: other central nervous system disease; history of severe head trauma, severe hypoglycemic, microvascular and macrovascular complications, alcohol dependence or poison use; any history of mental and psychological disease or family history; any hearing or visual impairment; contraindications for MRI examination. This study is examined and approved by the Medical Research Ethics Committee of Guangzhou University of Chinese Medicine. The demographic and clinical information of the two groups are summarized in Table 1. As illustrated in the table, the average duration of T2DM patients is 6.6 ($\pm$3.4) years.

2.2 MRI Acquisition

The MRI data are collected by a 3T scanner (GE Medical Systems, United States) with an 8-channel head coil. The conventional sequences which contain T1 weighted (reconstructed voxel size $1.00 \times 1.00 \times 1.00\ mm^3$), T2W weighted, and fluid attenuated inversion recovery (FLAIR) is conducted first for screening lesions, followed by the diffusion tensor imaging (DTI) scan. Participants are required to ensure enough sleep and avoid any needle injection before the scan. During the scans, the participants are instructed to avoid sleeping or head movement. The detailed scanning parameters are consistent with the previous work [10].

2.3 Data Preprocessing

The detailed MRI data preprocessing follows the previous works [10, 11] and utilizes the PANDA pipeline toolbox [12]. The diffusion MRI data are first corrected for eddy-current induced distortion and simple head-motion during scanning. The tensor matrix and the DT metrics of Fractional Anisotropy (FA) scalar are then calculated for each voxel. Based on the T1 image of each participant, the entire brain is parcellated using the automated anatomical labeling (AAL) atlas. Then we can identify 90 cortical and subcortical regions, where each region represents a node in the structural network (Figure 1). The constraint spherical deconvolution-based tractography is utilized to reconstruct white matter tracts. The whole-brain fiber tract reconstructions are then parcellated using the AAL atlas. Each connection is weighted by the FA because this measure has revealed to be a sensitive indicator of structural white matter disruptions in T2DM patients. These preprocessing steps generate a weighted connectivity matrix for each participant.

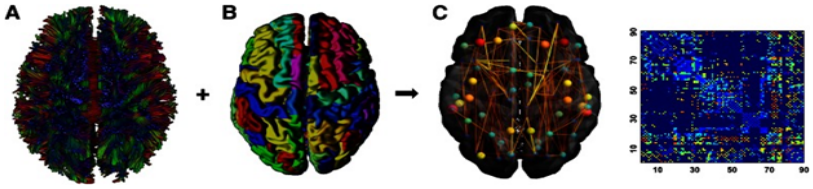

Figure 1: Structural brain network construction (This picture is adopted from previous work [12]).

2.4 Group Permutation Tests

As there are in total $90 \times 90 = 8100$ connections constructed for each participant, to briefly demonstrate the structural group information, we conducted permutation tests with 5000 randomizations for each connection to compare the FA strength between T2DM group and HC group following the common procedures [13] using the DPABI toolbox [14]. Multiple comparisons are corrected for with FDR correction ($p<0.025$). The connections with significant difference between the two groups are illustrated using Circos [15] and 3D figures, which generated by the BrainNet Viewer [16].

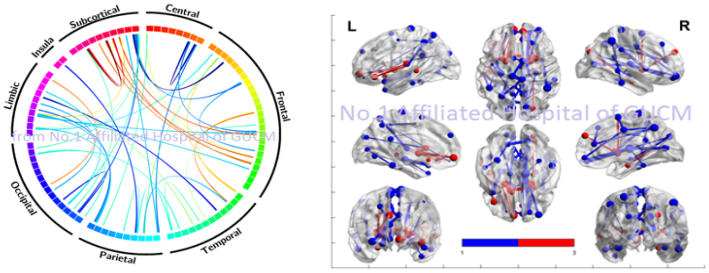

Figure 2: Structural connection with significant group difference.

3 Task 1: Distinguishing T2DM Patients and Healthy People

Our first task is to use the brain connectivity features for predicting whether the person is a T2DM patient. The most significant challenge here is that the number of connectivity features (8,100) is much larger than the sample size (94). Under this case, the fitted model has a high risk of being overfitting and shows poor performance in practical use. Therefore, it is important to adapt the feature selection techniques to screen out some features that contribute less to the model prediction.

To evaluate the fitted model's performance in practical use, we apply the stratified K-fold cross-validation strategy, i.e., the model is trained on (K-1) folds of data at a time, and the reserved one fold is used for validation. The training and evaluation process is performed K times. Our evaluation metrics include accuracy, precision, recall, and F1-score. Since each time of training and evaluation in the cross-validation can result in the set of metrics, the average and variance are considered.

3.1 Baseline Building

We first build some baseline models using logistic regression. The baseline models use the data: (1) All 8,100 brain connectivity features; (2) 90 surface size features; (3) 90 voxel size features; (4) 90 surface size features+90 voxel size features. The performance results of these baseline models are shown in Table 2. As can be seen from the results, the baseline models do not provide good performance for the classification task. However, the performance results still suggest that the connectivity between brain regions can provide more useful features for predicting T2DM patients compared to the voxel size and surface size of the brain regions.

	Accuracy	Precision	Recall	F1-score
(1)	0.5011	0.5009	0.5333	0.5104
	(0.0064)	(0.0085)	(0.0195)	(0.0105)
(2)	0.4678	0.4840	0.4733	0.4649
	(0.0056)	(0.0103)	(0.0161)	(0.0056)
(3)	0.3614	0.3478	0.3844	0.3610
	(0.0057)	(0.0111)	(0.0274)	(0.0179)
(4)	0.3824	0.3616	0.4044	0.3772
	(0.0066)	(0.0129)	(0.0376)	(0.0226)

Table 2: Performance for the baseline models (values in brackets: variances).

3.2 Feature Selection with Permutation Test

In order to build a better model for the prediction using the brain connectivity features, we try to make some feature selections before the model building. We start from the data side first. Here, we perform a permutation test for each brain connectivity feature to check whether the distribution of the feature is significantly different in the control and patient groups.

The basic idea of the permutation test is to test whether the difference in means between two groups is changed significantly when the group labels are randomly reassigned. The null hypothesis is that the data drawn from the control group is from the same distribution as the data drawn from the patient group. We reject the null hypothesis for each brain connectivity feature at the significance level of 0.05. Under the significance level, most of the connectivity do not show obvious differences between the control and patient groups, and the null hypothesis is rejected for only 47 of the 8,100 features.

We then build the classification model with the 47 features. The used models include logistic classification, ridge classification with the hyper-parameter α (controls the penalty to the model size) tuning, and support vector machine (SVM). The performance results for these models are shown in Table 3. The performance of these models is significantly better than the previous models with all connectivity features included. This indicates that by using the permutation test to filter out features that are not significantly different between the two groups, the preserved features are more discriminative for making the prediction of whether the person is a T2DM patient.

	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.7672	0.7756	0.7888	0.7727
	(0.0181)	(0.0278)	(0.023)	(0.0107)
Ridge Regression	0.7456	0.7794	0.7244	0.7396
(alpha=0.9, from tuning)	(0.0100)	(0.0242)	(0.0159)	(0.0103)
SVM	0.8093	0.8297	0.7867	0.8005
	(0.0137)	(0.0223)	(0.0238)	(0.0161)

Table 3: Performance for the models with features selected from permutation test (values in brackets: variances).

3.3 Forward Stepwise Feature Selection

To further screen out the redundant features that may lead to poor prediction and interpretation, we apply the forward stepwise selection strategy to look for the best subset for the classification. The AIC, BIC and Adjusted R^2 are reported for each subset with different amounts of predictors. The results are shown in Figure 3.

The three estimates of the test error indicate we need a model with 10-20 predictors. We test the models under these subsets and find that the models under the 18 selected features give the best performance in the cross-validation. The results are given in Table 3. We can see that compared to the baseline model (Table 2 (1)), the performance has been improved significantly with our carefully designed feature selection (via permutation test first, then apply forward stepwise feature selection).

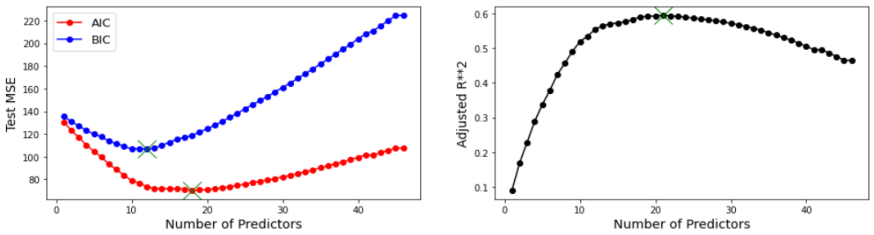

Figure 3: AIC, BIC and Adjusted R^2 during the forward feature selection.

	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.8099 (0.0071)	0.7983 (0.0072)	0.8311 (0.0162)	0.8104 (0.0083)
Ridge Regression (alpha=0.2, from tuning)	0.8304 (0.0058)	0.8694 (0.0066)	0.7867 (0.0189)	0.8185 (0.0089)
SVM	0.8093 (0.0026)	0.8127 (0.0035)	0.8088 (0.0106)	0.8065 (0.0038)

Table 4: Performance for the models with the 18 features from forward stepwise feature selection (values in brackets: variances).

3.4 Adding some Interaction Terms

In Sections 3.2 and 3.3, we explore how brain connectivity features can be selected to build models with better performance in practical use. In order to further improve the classification performance, we try to include some interaction terms in the classification model. The interaction term shows whether the effect of one predictor on the response variable depends on the values of another predictor. And some previous study has pointed out that the connectivity between different brain regions can be affected synergistically by the disease [3, 4, 7, 17].

We consider the second-order interaction, and the problem here is that the 18 selected features can yield 153 second-order interaction terms. Therefore, we need to select the terms that contribute significantly to the classification from these 153 interaction terms. A logistic regression function is then fitted with the interaction terms, and the forward stepwise selection method is applied. Figure 4 reports the AIC, BIC and adjusted R^2 for the best subset with different numbers of predictors (we restricted the maximum number of terms to 20). These metrics suggest that the models with 12-20 interaction terms can fit the data well. By using the original 18 connection features with different subsets of interaction terms (12-20) to build the model, we find that the model including 14 selected interaction terms gives the best classification performance (the results are given in Table 5).

With the introduction of the second-order interaction terms, the model performance is improved significantly, especially for the ridge regression model, whose accuracy is improved from 83.04% to 92.51%. The other metrics: precision, recall and F1-score are also outstanding compared to the previous models. The results confirm that the synergistic effects on the brain region connectivity by T2DM indeed exist.

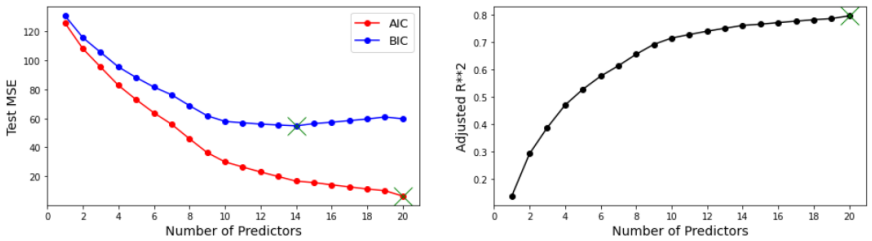

Figure 4: AIC, BIC and Adjusted R^2 during the forward feature selection of the interaction terms.

	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.8204 (0.0071)	0.8011 (0.0076)	0.8511 (0.0125)	0.8232 (0.0081)
Ridge Regression (alpha=0.3, from tuning)	0.9251 (0.0007)	0.9377 (0.0026)	0.9155 (0.0018)	0.9249 (0.0006)
SVM	0.8204 (0.0049)	0.8350 (0.0089)	0.8088 (0.0106)	0.8170 (0.0059)

Table 5: Performance for the Models with the 18 Features from Forward Stepwise Feature Selection (values in brackets: variances).

4 Task 2: Predicting the Disease Duration of Type-II Diabetes Patients

In the first task, we build an accurate classification model to distinguish between T2DM patients and healthy people. Next, we try to use the brain connectivity features to predict the time course of T2DM. The time course of the disease is an important reference factor for doctors in determining a patient's treatment plan. However, many patients often do not realize they have type-II diabetes in the early stages of the disease. It will be helpful for doctors if accurate duration predictions can be provided by using the brain connectivity features.

Since we have only 47 data instances for the T2DM patients, we apply the leave-one-out (LOO) cross-validation strategy to evaluate the performance of built models. The LOO strategy is well suited for use in small datasets like we have now. We use root mean square error (RMSE) as the evaluation metric, and the average and variance of the RMSE during the LOO cross-validation are considered.

4.1 Baseline Building

We first build some baseline models using the linear regression model. The baseline models use the data: (1) All 8,100 brain connectivity features; (2) 90 surface size features; (3) 90 voxel size features; (4) 90 surface size features+90 voxel size features. The performance results of these baseline models are shown in Table 6. As can be seen from the results, the connectivity between brain regions provides much more useful information for predicting the time course since the performance of the model built by the connectivity features is significantly better than models built by the surface size and voxel size. However, the performance of all these baseline models is not that satisfactory, especially the high variance of RMSE indicates the model prediction is unstable.

	(1)	(2)	(3)	(4)
RMSE	3.03	4.38	4.41	4.13
Var(RMSE)	5.35	12.20	8.95	8.67

Table 6: Performance for the Baseline Regression Models on the Disease Duration.

4.2 Data Transformation

Since the distribution of the disease duration in the dataset is far away from the normal distribution (with excess kurtosis=-0.739 and skewness=0.529) , the model fitted directly on the duration has high risk of breaking the normality assumption of the linear regression model. We do the square root transformation on the duration before the further model building, where $y^{(i)}_{transform} = \sqrt{y^{(i)}}$. In Section 4.4, we show that the regression model fitted with the transformed duration satisfies the linear regression model's assumptions.

4.3 Feature Selection

In order to build the model with better prediction performance, similar to what we do in the classification model, we use the 47 brain connectivity features selected from the permutation test and then adapt the forward stepwise feature selection. The AIC, BIC and Adjusted R^2 for each subset with different amounts of predictors are given in Figure 5.

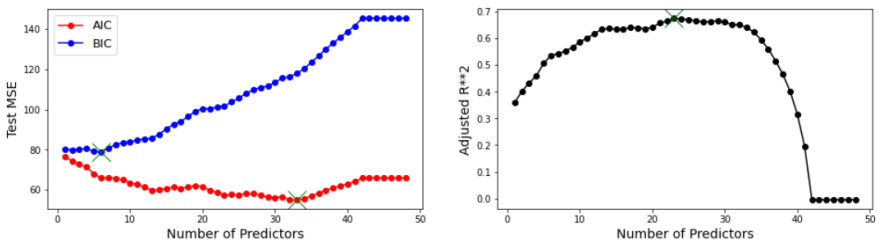

Figure 5: AIC, BIC and Adjusted R^2 during the forward feature selection for the regression model.

Although the AIC and Adjusted R^2 suggest a higher amount of features (32) can fit the data better, we choose to use the subset with six features as suggested by the BIC. Compared to AIC and Adjust R^2, BIC gives more penalty to the model size. Since we have a small dataset here, the model with high amount of features can be overfitted easily and result in poor performance in practical use. The linear regression model with the selected six brain connectivity features result in RMSE=1.94 and Variance(RMSE)=2.13 in the LOO cross-validation, where the prediction is much more stable now.

We further check whether the obtained regression function satisfies the five assumptions of linear regression, and the details are given in Appendix B. Since the fitted linear regression model with the selected six brain connectivity features satisfies all the five assumptions and has stable performance in the LOO cross-validation, we believe it is reliable for predicting the disease duration.

5 Discussion

As shown in Section 3 and 4, our experimental results yield quite promising performance in both distinguishing the diabetic brain and predicting the onset time of T2DM. In this section, We discuss on three major findings and future work suggestions: a) characterizing diabetic brain without the cognitive impairment; b) the progressively brain damage accompanied with the duration of T2DM; c) the advantage of connectivity features.

5.1 Diabetic Brain without the Cognitive Impairment

The best classification model of diabetic brain is developed from the 18 connection features which pass the permutation test and forward feature selection. The final model achieves quite promising accuracy in the cross validation (92.5%), which suggests that suffering T2DM can lead to distinguishable brain damaged that characterized the diabetic brain even before they show the cognitive impairment. Figure 6 show the selected 18 connections in the classification model. And the detailed information for the ridge regression model (include coefficients) is given in Appendix A. As can be seen, most of the connections are associated with a negative coefficient, suggesting that the fiber connection between the two regions of interest (ROI) are impaired in T2DM patients.

As observed, most of the 18 connections have already been reported in T2DM patients with cognitive impairment (e.g. connections to thalamus, visual regions, frontal lobes, and temporal lobes). For example, thalamus is reported to be a crucial candidate which response for the cognitive impairment in AD. Thalamus is the major gateway of sensory input and relays it to the temporal and frontal cortex [3]. The fronto-striato-thalamic circuits in type 2 diabetes is closely correlated with memory, execution function, learning and attention. Interestingly, we find that 3 out of 18 crucial connections in the model are related to the right thalamus (Figure 6). This suggests that the brain damage in critical regions related to memory and cognitive loss already occurs in those patients although they do not show clinical cognitive impairment yet. The structural scans seem to be more sensitive to the development of the diabetic brain than the clinical cognitive diagnosis. Therefore, it should be recommended for T2DM patients to conduct regular structural imaging scans to better capture their current status of their brain damage, which can provide them with earlier intervention if their brain connectivity information suggests potential MCI or even AD.

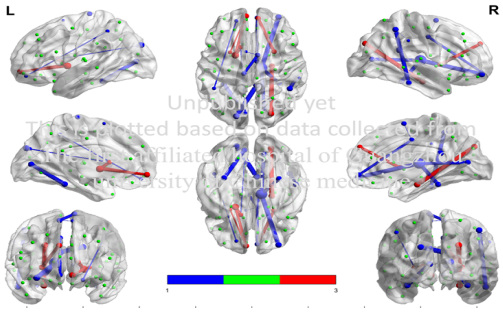

Figure 6: The 18 selected connections and the location of connected ROIs in the brain.

5.2 The Progressive Brain Damage

Another interesting finding is that we can predict the onset time of T2DM based on the information collected from current structural MRI scanning images with relatively small errors in the leave-one-out cross-validation ($\pm$1.9 years). The best model uses only six selected connections from the 47 connections which pass the permutation tests. This finding support that the brain shows progressive changes according to the duration of T2DM. For patients without any cognitive impairment (CI), it is still possible to detect and monitor their development of brain disruption during regular brain scans. Since the current study does not include patients with cognitive impairment, it is desirable to conduct future work to examine whether this model can be extended to T2DM patients with CI. This can help us to understand whether the progressive brain changes between CI and no-CI patients are different or not, which can further facilitate the brain damage prevention.

5.3 The Advantage of Connectivity Features

It is worth noticing that the structural connectivity features based on white matter fiber tracks outperform the regional features based on gray matter volume and surface size under all conditions. As reported in Table 2, the baseline model with connectivity features show better performance in all assessment metrics (Accuracy, Precision, Recall and F1-score). Since the baseline models do not penalize model size and connectivity features have more dimensions than regional features, we further explore how forward feature selection for regional features would improve its performance in classifying diabetic brains. Results are reported in Appendix C. As we can see, the best model using regional features is still worse than the selected model using connectivity features. Moreover, the connectivity features show less RMSE and more minor variance of LOO cross-validation performance in predicting T2DM duration. Those results support that the structural connections are more sensitive to hyperglycemia than gray matter atrophy measured by voxel-based methods. Since most previous works focus on the gray matter loss [2, 4, 5] and functional connectivity changes [17, 8], more studies focusing on structural connectivity and T2DM should be desirable.

6 Conclusion

In this project, we explore the disrupted brain connectivity in T2DM patients. We first try to build the classification model for distinguishing the diabetic brain from the healthy brain. The experimental results show that the diabetic brain becomes characterized before the patients show the measurable cognitive impairment since the classification model built based on the brain region connectivity features can show a high accuracy (92.5%). To further examine how the brain damage is developed progressively with the duration of T2DM, we build a regression model to use the connectivity features to predict the time course of T2DM patients. The final obtained linear regression model can predict onset of T2DM with the error $\pm$1.9 years. During the model building, we apply some feature selection techniques and find that structural connections based on white matter fiber tracks can better distinguish the diabetic brain than regional gray matter information, which has not been observed by previous works to the best of our knowledge.

References

[1] Haotian Qian, Dongxue Qin, Shouliang Qi, Yueyang Teng, Chen Li, Yudong Yao, and Jianlin Wu. Less is better: Single-digit brain functional connections predict t2dm and t2dm-induced cognitive impairment. *Frontiers in Neuroscience*, 14:1383, 2021.

[2] Fang Fang, Ya-Feng Zhan, Yao-Yao Zhuo, Da-Zhi Yin, Kang-An Li, and Yu-Fan Wang. Brain atrophy in middle-aged subjects with type 2 diabetes mellitus, with and without microvascular complications. *Journal of diabetes*, 10(8):625–632, 2018.

[3] Hao Lei, Rong Hu, Guanghua Luo, Tingqian Yang, Hui Shen, Hao Deng, Chunyu Chen, Heng Zhao, and Jincai Liu. Altered structural and functional mri connectivity in type 2 diabetes mellitus related cognitive impairment: A review. *Frontiers in Human Neuroscience*, 15, 2021.

[4] Chris Moran, Thanh G Phan, Jian Chen, Leigh Blizzard, Richard Beare, Alison Venn, Gerald Münch, Amanda G Wood, Josephine Forbes, Timothy M Greenaway, et al. Brain atrophy in type 2 diabetes: regional distribution and influence on cognition. *Diabetes care*, 36(12):4036–4042, 2013.

[5] Mingrui Li, Yifan Li, Yujie Liu, Haoming Huang, Xi Leng, Yuna Chen, Yue Feng, Xiaomeng Ma, Xin Tan, Yi Liang, et al. Altered hippocampal subfields volumes is associated with memory function in type 2 diabetes mellitus. *Frontiers in neurology*, 12, 2021.

[6] Li Yao, Chengmin Yang, Wenjing Zhang, Siyi Li, Qian Li, Lizhou Chen, Su Lui, Graham J Kemp, Bharat B Biswal, Nadim J Shah, et al. A multimodal meta-analysis of regional structural and functional brain alterations in type 2 diabetes. *Frontiers in neuroendocrinology*, 62:100915, 2021.

[7] Yuna Chen, Yongsheng Pan, Junshen Lu, Xin Tan, Yi Liang, Yifan Li, Wenjiao Lyu, Chunhong Qin, Haoming Huang, Shangyu Kang, et al. Identifying type 2 diabetic brains by investigating disease-related structural changes in mri. *Frontiers in Neuroscience*, page 1418, 2021.

[8] Renping Yu, Han Zhang, Le An, Xiaobo Chen, Zhihui Wei, and Dinggang Shen. Connectivity strength-weighted sparse group representation-based brain network construction for m ci classification. *Human brain mapping*, 38(5):2370–2383, 2017.

[9] Wen-bin Guo, Feng Liu, Jin-dong Chen, Xi-jia Xu, Ren-rong Wu, Chao-qiong Ma, Keming Gao, Chang-lian Tan, Xue-li Sun, Chang-qing Xiao, et al. Altered white matter integrity of forebrain in treatment-resistant depression: a diffusion tensor imaging study with tract-based spatial statistics. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 38(2):201–206, 2012.

[10] Ying Xiong, Tian Tian, Yang Fan, Shaolin Yang, Xiaoxiao Xiong, Qiang Zhang, and Wenzhen Zhu. Diffusion tensor imaging reveals altered topological efficiency of structural networks in type-2 diabetes patients with and without mild cognitive impairment. *Journal of Magnetic Resonance Imaging*, 55(3):917–927, 2022.

[11] Yael D Reijmer, Alexander Leemans, Manon Brundel, L Jaap Kappelle, Geert Jan Biessels, and Utrecht Vascular Cognitive Impairment (VCI) Study Group. Disruption of the cerebral white matter network is related to slowing of information processing speed in patients with type 2 diabetes. *Diabetes*, 62(6):2112–2115, 2013.

[12] Zaixu Cui, Suyu Zhong, Pengfei Xu, Yong He, and Gaolang Gong. Panda: a pipeline toolbox for analyzing brain diffusion images. *Frontiers in human neuroscience*, 7:42, 2013.

[13] Anderson M Winkler, Gerard R Ridgway, Gwenaëlle Douaud, Thomas E Nichols, and Stephen M Smith. Faster permutation inference in brain imaging. *Neuroimage*, 141:502–516, 2016.

[14] Chao-Gan Yan, Xin-Di Wang, Xi-Nian Zuo, and Yu-Feng Zang. Dpabi: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, 14(3):339–351, 2016.

[15] Martin Krzywinski, Jacqueline Schein, Inanc Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J Jones, and Marco A Marra. Circos: an information aesthetic for comparative genomics. *Genome research*, 19(9):1639–1645, 2009.

[16] Mingrui Xia, Jinhui Wang, and Yong He. Brainnet viewer: a network visualization tool for human brain connectomics. *PloS one*, 8(7):e68910, 2013.

[17] Zhenyu Liu, Jiangang Liu, Huijuan Yuan, Taiyuan Liu, Xingwei Cui, Zhenchao Tang, Yang Du, Meiyun Wang, Yusong Lin, and Jie Tian. Identification of cognitive dysfunction in patients with t2dm using whole brain functional connectivity. *Genomics, proteomics & bioinformatics*, 17(4):441–452, 2019.

[18] Theodore W Anderson and Donald A Darling. Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, pages 193–212, 1952.

[19] James Durbin and Geoffrey S Watson. Testing for serial correlation in least squares regression. iii. *Biometrika*, 58(1):1–19, 1971.

Appendix

A. Detailed Information for the Classification Models

	Coefficients	Standardized Coefficient	Standard Errors	t values	Probabilities
intercept	3.0459	NA	0.543	5.605	0.000
['Caudate_L', 'Pallidum_L']	0.6747	0.066316	0.172	3.917	0.000
['Fusiform_R', 'Thalamus_R']	-0.4652	-0.062949	0.120	-3.862	0.000
['Frontal_Med_Orb_R', 'Cingulum_Mid_R']	-1.1464	-0.093789	0.196	-5.859	0.000
['Cingulum_Post_R', 'Occipital_Sup_L']	-0.3565	-0.057629	0.105	-3.402	0.001
['Calcarine_R', 'Thalamus_R']	-1.3564	-0.141233	0.166	-8.159	0.000
['Frontal_Mid_Orb_R', 'Thalamus_R']	-1.5043	-0.188765	0.137	-10.987	0.000
['Frontal_Sup_Orb_L', 'Parietal_Inf_L']	-1.7015	-0.094766	0.284	-5.996	0.000
['Supp_Motor_Area_R', 'Paracentral_Lobule_L']	-1.3261	-0.117846	0.198	-6.712	0.000
['Cuneus_L', 'Caudate_L']	-2.0824	-0.107192	0.331	-6.295	0.000
['Parietal_Inf_R', 'Temporal_Inf_R']	-0.9017	-0.117828	0.127	-7.124	0.000
['ParaHippocampal_R', 'Occipital_Sup_R']	0.8107	0.103384	0.130	6.216	0.000
['Calcarine_L', 'Fusiform_L']	-0.7449	-0.092779	0.130	-5.739	0.000
['Parietal_Inf_R', 'SupraMarginal_R']	-3.3436	-0.055836	0.988	-3.384	0.001
['Supp_Motor_Area_R', 'Cingulum_Mid_R']	-3.1377	-0.054535	0.933	-3.363	0.001
['Supp_Motor_Area_L', 'Supp_Motor_Area_R']	1.5698	0.059334	0.457	3.434	0.001
['Frontal_Med_Orb_L', 'Pallidum_L']	0.8031	0.084796	0.164	4.888	0.000
['Frontal_Med_Orb_L', 'Putamen_L']	-0.5418	-0.067832	0.145	-3.742	0.000
['Frontal_Sup_Medial_R', 'Putamen_R']	0.5617	0.065531	0.146	3.858	0.000

Table 7: Ridge Regression Model with the 18 Selected Brain Connectivity Features (for classification).

B. Testing Linear Regression Assumptions

We check that whether the obtained linear regression function satisfies the five assumption of the linear regression model: (1) Linear relationship; (2) Multivariate normality; (3) No or little multicollinearity; (4) No auto-correlation; (5) Homoscedasticity.

(1) We plot the graph of predicted values against the ground truth values (Figure 7). Since the points are evenly distributed around the diagonal line, the linearity assumption is satisfied;

(2) We plot the distribution of the residuals (Figure 8) and also use the Anderson-Darling test [18] to check the normality of the residuals. The plot and p-value (0.3659) from the test confirm that the normality of the residuals for the fitted model is acceptable;

(3) The correlation between different features are checked. As the highest absolute correlation value is 0.12, the assumption of small multicollinearity is satisfied;

(4) We used the Durbin-Watson test [19] to check whether the error terms are autocorrelated. The resultant test value is 2.19, indicating that there is almost no autocorrelation;

(5) We plot the residuals (Figure 9) and find that the variance of the residuals appears to be uniform. Therefore, the homoscedasticity is considered to be satisfied.

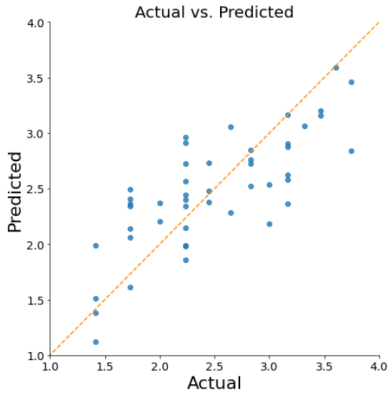

Figure 7: Graph of predicted values versus the ground truth values for the regression model.

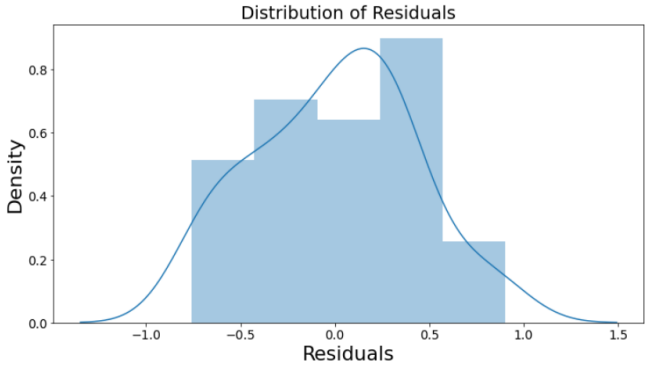

Figure 8: Distribution of the residuals for the regression model.

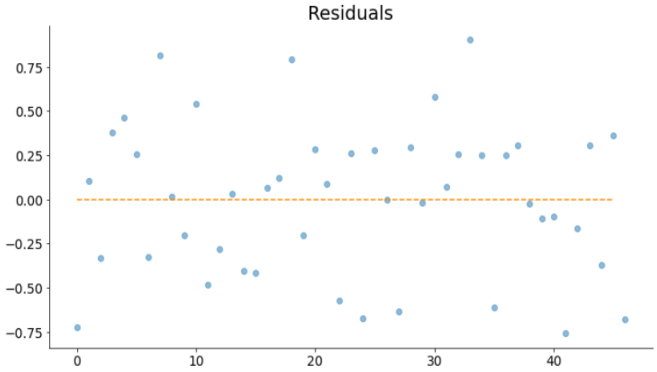

Figure 9: Scatter plot of the residuals for the regression model.

C. Feature Selection for Classification Models via Surface Size and Voxel Size

C.1 Surface Size with Feature Selection

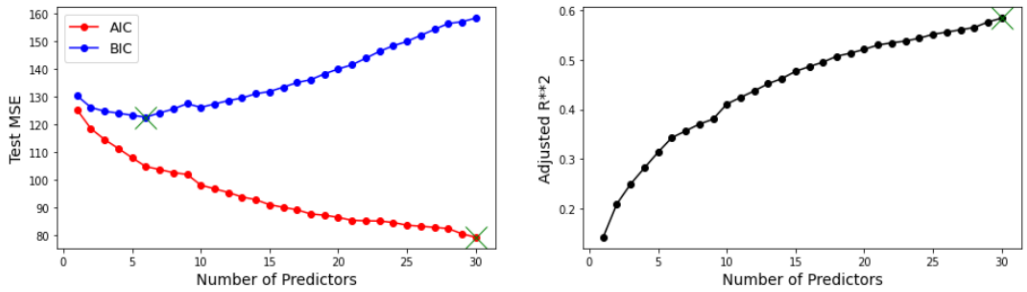

Figure 10: AIC, BIC and Adjusted R^2 during the forward feature selection for the surface size.

	Accuracy	Precision	Recall	F1-score
N=6	0.7766 (0.0026)	0.8109 (0.0100)	0.7733 (0.0360)	0.7627 (0.0094)
N=30	0.7860 (0.0063)	0.7915 (0.0091)	0.8111 (0.0131)	0.7914 (0.0051)
N=90	0.4678 (0.0057)	0.4841 (0.0103)	0.4733 (0.0161)	0.4649 (0.0056)

Table 8: Performance for the models with different amount of surface size features (values in brackets: variances).

C.2 Voxel Size with Feature Selection

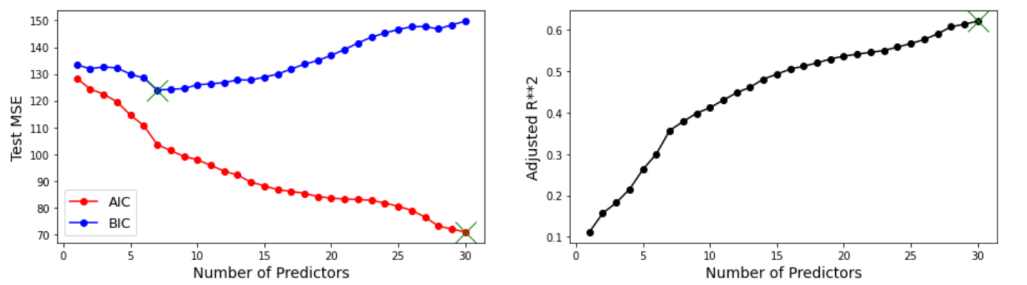

Figure 11: AIC, BIC and Adjusted R^2 during the forward feature selection for the voxel size.

	Accuracy	Precision	Recall	F1-score
N=6	0.7766 (0.0015)	0.7694 (0.0045)	0.8089 (0.0106)	0.7812 (0.0019)
N=30	0.6801 (0.0127)	0.6950 (0.0262)	0.6978 (0.0086)	0.6882 (0.0111)
N=90	0.3614 (0.0057)	0.3478 (0.0111)	0.3844 (0.0274)	0.3610 (0.0179)

Table 9: Performance for the models with different amount of voxel size features (values in brackets: variances).

Statement of Contribution

Yue Wei:

1. Data acquisition and preprocessing;
2. Building of regression models;
3. Writing report and preparing PowerPoint;
4. Making presentation.

Weiyan Xie:

1. Building of classification and regression models;
2. Writing report and preparing PowerPoint;
3. Making presentation.