Pairwise Learning for Autism Spectrum Disorder Imbalanced Classification

Shu Liu
Computational and Data Science PhD Program
Middle Tennessee State University
Murfreesboro, TN, USA
sl6b@mtmail.mtsu.edu

Qiang Wu
Department of Mathematics
University of Tennessee
Knoxville, TN, USA
Qiang.Wu@utk.edu

Xin Yang
Department of Computer Science
Middle Tennessee State University
Murfreesboro, TN, USA
Xin.Yang@mtsu.edu

2024 IEEE 6th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS) | 979-8-3503-9613-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ECBIOS61468.2024.10885487

***Abstract*—This paper performs a classification task on data obtained from the Autism Brain Imaging Data Exchange (ABIDE) repository. In real-world case analysis, the number of autism spectrum disorder (ASD) patients is much smaller than typically developed people. To address this issue, this paper proposes the utilization of pairwise robust support vector machine (PRSVM) algorithms to classify autism spectrum disorder (ASD) patients. In this project's experiments, the correlation matrix derived from functional magnetic resonance imaging (fMRI) data was employed as a classification feature. A comprehensive evaluation was conducted to compare the classification performance of PRSVM with various machine learning methods. The comparative analysis encompassed various aspects, including different data dimensions, imbalanced ratios, and sample sizes, providing valuable insights into the relative performance of the algorithms under different experimental conditions. The experimental results demonstrate that PRSVM can detect autistic patients more accurately when the data is imbalanced. Moreover, the results indicate that PRSVM outperforms or achieves comparable performance to other conventional classification methods in a variety of situations. Furthermore, our approach can be further improved by augmenting the training set with either exclusively normal person samples or by incorporating patient samples and normal people samples in a proportionate manner. This augmentation strategy holds promising application value, as it contributes to improving the performance and robustness of our method.**

Keywords— fMRI, ASD classification, pairwise robust support vector machine, imbalanced data

I. Introduction

A. Functional magnetic resonance imaging

functional magnetic resonance imaging (fMRI) is a non-invasive neuroimaging technique that measures neuronal activity by detecting changes in blood oxygenation level dependent (BOLD) signal. This technique is based on the principle that changes in neural activity within the brain are accompanied by corresponding changes in local blood flow and oxygenation level. The BOLD signal is derived from the differences in magnetic properties between oxygenated and deoxygenated blood. fMRI provides a way to indirectly measure brain activity and has emerged as a prevalent technique to investigate functional connectivity, brain networks, and activation patterns [1].

In 1980, Roy and Sherrington [2] found that regional cerebral blood flow could serve as an indicator of neuronal viability in the corresponding brain area. The pioneering work by Ogawa et al. [1] in 1990 introduced the concept of BOLD, which subsequently enabled the realization of the fMRI imaging technique. In 1991, researchers achieved a groundbreaking milestone by demonstrating the first-ever visualization of both brain structure and function using fMRI [3]. The underlying principle for fMRI brain imaging is that the increase of local neuronal activity often leads to increased oxygen demand. The oxygen in oxyhemoglobin will produce a paramagnetic molecule called deoxyhemoglobin. During the examination, accumulated deoxyhemoglobin can act as a local contrast agent to enhance local signal intensity. Thus, natural contrast agents can target task-relevant brain regions and visualize them with fMRI.

fMRI has a profound impact on the field of cognitive neuroscience. Since its discovery in 1990, fMRI has rapidly developed into one of the most commonly used techniques in the discipline. Especially in the field of treatment of mental illness, fMRI has become an auxiliary diagnostic tool. So far, fMRI has been used to discover the abnormal brain functionality associated with a wide range of mental diseases [4]-[6]. In traditional classification studies involving fMRI data, the most commonly utilized machine learning techniques are support vector machines (SVM) and kernel SVM [7],[8]. SVM is a well-known supervised learning algorithm for handling high-dimensional data and has proven to be particularly effective in the realm of fMRI analysis. Kernel SVM extends the capabilities of SVM by employing various types of kernel functions, such as linear, polynomial, and radial basis function (RBF), which can capture non-linear relationships and improve classification performance.

However, in real-life scenarios, the number of patients with a condition of interest, such as ASD, is often much smaller than the number of normal people. Dealing with imbalanced fMRI datasets poses a challenge for traditional classification methods, as they always tend to overfocus on the majority class. As a result, identifying an effective classifier for imbalanced fMRI datasets has become an important research area for scholars. In recent years, a variety of machine learning algorithms have been used to deal with imbalanced fMRI data classification [9],[10], these algorithms are on the basis of oversampling and synthetic minority oversampling technique, which generates some samples that are not informative and increased likelihood of overfitting. To address this issue, this paper proposes the utilization of pairwise robust support vector machine (PRSVM) algorithms to classify ASD patients.

B. Autism spectrum disorder classification

Autism was discovered and named by Kanner [11] in 1943, and it is recognized as a special type of developmental disorder by the World Health Organization and the American Psychiatric Association. The current consensus is that deficits in social and verbal communication skills and repetitive stereotyped behaviors manifested before the age of three are the defining characteristics of children with autism.

Autism Brain Imaging Data Exchange (ABIDE) [12] is a data-sharing initiative designed to advance research on autism spectrum disorder (ASD). The project collects and shares neuroimaging data from multiple authoritative institutions, and these datasets contain multiple modalities of data including structural MRI, functional MRI, magnetic resonance spectroscopy, and magnetic resonance diffusion tensor imaging. This initiative currently has two large-scale collections: ABIDE I and ABIDE II, of which ABIDE collected 1112 datasets, of which 539 were from autistic patients, 573 were from typical controls, and ABIDE II collected 1114 datasets, of which 521 from autistic patients, 593 from typical controls.

With the open sharing of ABIDE data, many analyzes of ABIDE have emerged, and various machine learning algorithms such as empirical bayes, logistic regression, and SVM [13]-[15] are used to classify ASD patients. The experimental results of many articles show that when all samples are used for classification, machine learning algorithms usually only achieve an accuracy rate of 60% to 70% [14]. The machine learning algorithm can only achieve reliable accuracy when the total sample size is less than 100 [16]. With the rise and development of neural networks, more deep learning algorithms are used to improve the accuracy of ABIDE data classification [17]-[19].

In this paper, we use the pairwise robust support vector machine [20] to explore the classification accuracy when the classification labels are imbalanced for the data from ABIDE I. In essence, our aim is to identify patients with ASD in a large number of typical controls. The goal is to employ robust and accurate classification models that can mitigate the inherent imbalance in the data.

II. Methodology

A. Principal component analysis

We know that if there is a strong linear correlation between certain dimensions in the data, the information provided by the sample on these two dimensions will be repeated to a certain extent. So we hope that the dimensions of the input are orthogonal. In addition, the dimension of the correlation matrix is too large. In order to reduce the calculation amount of data processing, we choose to use principal component analysis (PCA) to reduce the dimensionality of the data.

Supposed a data of n observations $\mathrm{X} = \{x_1, x_2, \dots, x_\mathrm{n}\}$ with $x_\mathrm{i} \in \mathbb{R}^p$, the original data can be regarded as a matrix with n rows and p columns. Assume that the mean of each dimension of the original data is 0, we let this matrix multiply an $p * p$ orthogonal transformation matrix W, where W consists of column vectors $\{\mathrm{W}_1, \mathrm{W}_2, \dots, \mathrm{W}_p\}$. Then the original data is transformed into a new coordinate system. To fix the values of the data, each column vector $\|\mathrm{W_i}\| = 1$, the matrix after dimensionality reduction is $\mathrm{T} = \mathrm{XW}$, each column vector in T is $\{\mathrm{t}_1, \mathrm{t}_2, \dots, \mathrm{t}_p\}$. In order to compute the transformation matrix W, we need to compute the eigenvalues and eigenvectors of the covariance matrix $\mathrm{C} = \frac{1}{p}\mathrm{XX}^\mathrm{T}$,

the eigenvectors could be combined into a change matrix W

from left to right in the order of eigenvalues from large to small. Here, we can only keep the eigenvectors with big eigenvalues to reduce the dimensionality. If W' is used to represent the change matrix after discarding the eigenvectors with smaller eigenvalues, where W' consists of column vectors $\{\mathrm{W}_1', \mathrm{W}_2', \dots, \mathrm{W}_\mathrm{k}'\}$ and $\mathrm{k} < \mathrm{p}$. $\mathrm{T}' = \mathrm{XW}'$ is the data after dimensionality reduction to k dimension.

B. Pairwise robust support vector machine

For Pairwise robust support vector machine (PRSVM) algorithm handles the task of imbalanced data classification by using the robust support vector classifiers (RSVC) loss [21]

$$\mathrm{L_{RSVC}}\big(yf(x)\big) = \sigma^2\left(1 - \exp\left(\frac{\big(1 - yf(x)\big)_+^2}{\sigma^2}\right)\right)$$

in a pairwise learning framework. Here, σ is a tunable parameter.

We assume that the label of autistic patients is 1 and the label of typical controls is -1. In the classification of autism spectrum disorder, the binary classifier f should provide with a result of $\mathrm{sign}\left(f\left(x_{\text{autistic patient}}\right)\right) = 1$ and $\mathrm{sign}\left(f\left(x_{\text{typical control}}\right)\right) = -1$. For each pair of observations $(x_\mathrm{i}, y_\mathrm{i})$and $\left(x_\mathrm{j}, y_\mathrm{j}\right)$, if x_i represents an autistic patient and x_j represents a typical control, a good real-valued f will produce the result such that $\mathrm{f}(x_\mathrm{i}) - \mathrm{f}\big(x_\mathrm{j}\big) > 0$; on the contrary, if x_i represents a typical control and x_j represents an autistic patient, it should be $\mathrm{f}(x_\mathrm{i}) - \mathrm{f}\big(x_\mathrm{j}\big) < 0$; when both x_i

and x_j represent an autistic patient or a typical control simultaneously, we have no expectations. We let $y_\mathrm{ij} = \big(y_\mathrm{i} - y_\mathrm{j}\big)/2$ and write up the notation

$$y_{ij} = \begin{cases} 1, & if\ y_i = 1\ and\ y_j = -1; \\ -1, & if\ y_i = -1\ and\ y_j = 1, \end{cases}$$

if the algorithm has good performance, we can image that we will get $y_\mathrm{ij}\left(f(x_\mathrm{i}) - f\big(x_\mathrm{j}\big)\right) > 0$.

The pairwise loss could be written as

$$\mathrm{L}\Big(f, (x_i, y_i), \big(x_j, y_j\big)\Big) =$$

$$\sigma^2\left(1 - \exp\left(\frac{\Big(1 - y_{ij}\big(f(x_i) - f\big(x_j\big)\big)\Big)_+^2}{\sigma^2}\right)\right).$$

This loss is calculated by pairing a sample from the minority class with a sample from the majority class, ensuring that both classes make equal contributions during model training.

To compare with support vector machines, here we focus on linear classifiers $\mathrm{f}(x) = w^\mathrm{T}\mathrm{x} + \mathrm{b}$ with $\mathrm{w} \in \mathbb{R}^\mathrm{p}$ and $\mathrm{b} \in \mathbb{R}$.Then the optimization problem is

$$\min_w \sum_{i=1}^{n} \sum_{j:y_j \neq y_i} \sigma^2\left(1 - \exp\left(\frac{\Big(1 - y_{ij}w^\mathrm{T}\big(x_i - x_j\big)\Big)_+^2}{\sigma^2}\right)\right).$$

For $b \in \mathbb{R}$, let $\mathcal{E}_+(b)$ represent the false positive rate (FPR), and $\mathcal{E}_-(b)$ represent the false negative rate (FNR) of the classifier $\widehat{w}^\top x + b$. Let's define

$$\mathcal{E}(b) = max\big(\mathcal{E}_+(b), \mathcal{E}_-(b)\big)$$

and the intercept is estimated by

$$\hat{b} = \min_{b\in\mathbb{R}} \mathcal{E}(b)$$

When the sample sizes of autistic patients and typical controls are imbalanced, PRSVM combines the observations of the minority class and the majority class into pairs and then enters the model. This method can effectively balance the influence of the two classes.

C. Experiments and results

In this research, the data sets are extracted from two regions-of-interests (ROI) atlas: Harvard-Oxford (HO) [22] and Eickhoff-Zilles (EZ) [23]. In the experiments of this paper, we set the label of autistic patients to 1 and typical controls label to -1.

SVM, AdaBoost and Naive Bayes (NB) are used for comparison with PRSVM. PRSVM demonstrates low sensitivity to the choice of the parameter σ, hence σ was set to 1 across all experiments. All other methods are implemented in R using standard packages. Specifically, for SVM, the train function from the caret package was utilized. To ensure a fair comparison, the method svmlinear was employed to generate linear classifiers, with all other parameters maintained at their default settings. For AdaBoost, the boosting function from the adabag package was used. The number of iterations is fixed at 200. Naive Bayes was implemented using the naiveBayes function from the e1071 package with default parameters.

We will evaluate the FPR and FNR simultaneously. A balanced FPR and FNR imply that the minority class has been equally addressed. Additionally, we evaluate the area under the receiver operating characteristic (ROC) curve (AUC), which is widely regarded as a balanced accuracy metric for imbalanced data classification problems. We randomly sampled $1/2$ of positive cases and $1/2$ of negative cases to create a training set, while the remaining cases were used as the test set. The number of repetitions for each experiment was 20. All reported results are the average of these 20 repetitions.

1) Experiment I: data with varying input dimensions

In the first experiment, we randomly sampled 100 samples from the autistic patients and 500 samples from the typical controls, resulting in an imbalance ratio of 1:5. The PCA method was used to reduce the dimensionality, and we retained 10 and 50 principal components, respectively. The first 10 principal components could achieve greater than 90% cumulative contribution rate while the first 50 principal components could achieve greater than 95%. The results are shown in Table I and Table II.

The results shows that when the imbalance ratio reaches 1:5, the average AUC of SVM, AdaBoost and Naive Bayes are close to 0.5, which means these models are ineffective in this case. PRSVM has the best performance on both 10-dimensional and 50-dimensional datasets. For some fixed sampled datasets, we achieve AUC close to 0.8 and accruary close to 80%.

TABLE I. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-HO** DATA SET WITH VARYING PRINCIPAL COMPONENTS

Input dimension	Method	FNR	FPR	AUC
10	PRSVM	0.4440 (0.0125)	0.4205 (0.0105)	0.6037 (0.0048)
	SVM	0.9120 (0.0075)	0.0455 (0.0048)	0.5213 (0.0020)
	AdaBoost	0.7060 (0.0063)	0.2095 (0.0049)	0.5422 (0.0035)
	NB	0.7440 (0.0115)	0.1640 (0.0095)	0.5460 (0.0030)
50	PRSVM	0.5610 (0.0093)	0.2890 (0.0085)	0.6259 (0.0050)
	SVM	0.6780 (0.0076)	0.1805 (0.0058)	0.5708 (0.0035)
	AdaBoost	0.7590 (0.0112)	0.1405 (0.0061)	0.5503 (0.0038)
	NB	0.7010 (0.0199)	0.1775 (0.0141)	0.5608 (0.0039)

TABLE II. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-EZ** DATA SET WITH VARYING PRINCIPAL COMPONENTS

Input dimension	Method	FNR	FPR	AUC
10	PRSVM	0.4070 (0.0097)	0.4920 (0.0095)	0.5831 (0.0045)
	SVM	0.9520 (0.0051)	0.0295 (0.0031)	0.5093 (0.0014)
	AdaBoost	0.7000 (0.0070)	0.2315 (0.0042)	0.5343 (0.0036)
	NB	0.7490 (0.0139)	0.1715 (0.0077)	0.5398 (0.0040)
50	PRSVM	0.5640 (0.0087)	0.2800 (0.0062)	0.6131 (0.0050)
	SVM	0.6560 (0.0082)	0.1895 (0.0047)	0.5728 (0.0038)
	AdaBoost	0.7780 (0.0089)	0.1465 (0.0063)	0.5378 (0.0032)
	NB	0.7100 (0.0197)	0.1890 (0.0144)	0.5505 (0.0037)

2) Experiment II: data with varying imbalance ratios

In real-world applications, collecting a large number of patient data for a medical institution is difficult. Here, we would like to investigate if our model could perform better by just adding samples from typical controls to its training data. We randomly sampled 100 from the autistic patients and 100, 250, and 500 from the typical controls. In this experiment, we retained 10 principal components. The results are shown in Table III and IV.

TABLE III. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-HO** DATA SET WITH VARYING IMBALANCE RATIOS

Input dimension	Method	FNR	FPR	AUC
1:1	PRSVM	0.4110 (0.0161)	0.4390 (0.0192)	0.5969 (0.0066)
	SVM	0.3930 (0.0095)	0.4420 (0.0088)	0.5825 (0.0039)
	AdaBoost	0.4700 (0.0087)	0.4320 (0.0085)	0.5510 (0.0043)
	NB	0.4900 (0.0199)	0.4110 (0.0195)	0.5505 (0.0052)
1:2.5	PRSVM	0.4660 (0.0141)	0.3936 (0.0136)	0.5892 (0.0035)
	SVM	0.9830 (0.0026)	0.0076 (0.0012)	0.5047 (0.0012)
	AdaBoost	0.7770 (0.0061)	0.1592 (0.0036)	0.5319 (0.0027)
	NB	0.8390 (0.0062)	0.1040 (0.0047)	0.5285 (0.0023)

1:5	PRSVM	0.5080 (0.0098)	0.3470 (0.0075)	0.6058 (0.0027)
	SVM	1.0000 (0.0000)	0.0000 (0.0000)	0.5000 (0.0000)
	AdaBoost	0.9230 (0.0020)	0.0480 (0.0011)	0.5145 (0.0009)
	NB	0.9420 (0.0026)	0.0200 (0.0008)	0.5190 (0.0012)

TABLE IV. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-EZ** DATA SET WITH VARYING IMBALANCE RATIOS

Input dimension	Method	FNR	FPR	AUC
1:1	PRSVM	0.4740 (0.0156)	0.4540 (0.0128)	0.5596 (0.0059)
	SVM	0.4040 (0.0104)	0.4870 (0.0102)	0.5545 (0.0036)
	AdaBoost	0.4810 (0.0095)	0.4940 (0.0099)	0.5195 (0.0042)
	NB	0.4420 (0.0162)	0.4970 (0.0174)	0.5305 (0.0035)
1:2.5	PRSVM	0.5390 (0.0127)	0.3660 (0.0090)	0.5688 (0.0042)
	SVM	0.9850 (0.0031)	0.0072 (0.0016)	0.5039 (0.0008)
	AdaBoost	0.7770 (0.0064)	0.1672 (0.0039)	0.5279 (0.0032)
	NB	0.8830 (0.0057)	0.0976 (0.0047)	0.5097 (0.0025)
1:5	PRSVM	0.4700 (0.0093)	0.4078 (0.0089)	0.6061 (0.0027)
	SVM	1.0000 (0.0000)	0.0000 (0.0000)	0.5000 (0.0000)
	AdaBoost	0.9470 (0.0023)	0.0446 (0.0012)	0.5042 (0.0009)
	NB	0.9720 (0.0021)	0.0148 (0.0008)	0.5066 (0.0008)

The results of Experiment II demonstrate that when investigating the impact of an imbalance ratio, ranging from 1:1 to 1:2.5 and further to 1:5, the FPR and FNR of PRSVM showed minimal variation compared to other methods with poor performance. Additionally, the AUC of PRSVM exhibited a slight increase.

3) Experiment III: data with varying sample size

We are interested in investigating whether we can enhance the model's performance by augmenting the number of samples in the training set while preserving a constant imbalance ratio.

In the third experiment, we increased the sample size proportionally while maintaining a fixed imbalance ratio of 1:2. Specifically, we randomly sampled 50, 150, and 250 data from the positive class, and 100, 300, and 500 data from the negative class, respectively. In this experiment, we retained 10 principal components. The results are shown in Table V and Table VI.

TABLE V. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-HO** DATA SET WITH VARYING SAMPLE SIZES

Input dimension	Method	FNR	FPR	AUC
25:50	PRSVM	0.4920 (0.0215)	0.4540 (0.0207)	0.5674 (0.0091)
	SVM	0.8580 (0.0129)	0.0740 (0.0055)	0.5340 (0.0048)
	AdaBoost	0.7020 (0.0144)	0.2310 (0.0070)	0.5335 (0.0074)
	NB	0.7600 (0.0173)	0.1780 (0.0163)	0.5310 (0.0044)
75:150	PRSVM	0.4730 (0.0116)	0.3687 (0.0090)	0.6195 (0.0023)
	SVM	0.9413 (0.0049)	0.0300 (0.0026)	0.5143 (0.0013)
	AdaBoost	0.7147 (0.0038)	0.2147 (0.0028)	0.5353 (0.0022)
	NB	0.7967 (0.0077)	0.1190 (0.0047)	0.5422 (0.0021)
125:250	PRSVM	0.3504 (0.0047)	0.4522 (0.0040)	0.6348 (0.0018)
	SVM	0.9820 (0.0020)	0.0096 (0.0011)	0.5042 (0.0005)
	AdaBoost	0.7112 (0.0028)	0.1984 (0.0022)	0.5452 (0.0012)
	NB	0.7690 (0.0026)	0.1136 (0.0020)	0.5452 (0.0009)

TABLE VI. CLASSIFICATION PERFORMANCE OF FOUR CLASSIFIERS ON **ROI-EZ** DATA SET WITH VARYING SAMPLE SIZES

Input dimension	Method	FNR	FPR	AUC
25:50	PRSVM	0.5000 (0.0239)	0.4110 (0.0226)	0.5517 (0.0092)
	SVM	0.9000 (0.0106)	0.0560 (0.0057)	0.5220 (0.0038)
	AdaBoost	0.7020 (0.0180)	0.2490 (0.0110)	0.5245 (0.0067)
	NB	0.7960 (0.0174)	0.1640 (0.0099)	0.5200 (0.0078)
75:150	PRSVM	0.4467 (0.0096)	0.4316 (0.0085)	0.5963 (0.0021)
	SVM	0.9653 (0.0040)	0.0163 (0.0018)	0.5092 (0.0013)
	AdaBoost	0.6960 (0.0044)	0.2233 (0.0025)	0.5403 (0.0027)
	NB	0.7540 (0.0070)	0.1667 (0.0053)	0.5397 (0.0018)
125:250	PRSVM	0.4392 (0.0067)	0.4040 (0.0069)	0.6093 (0.0016)
	SVM	0.9888 (0.0018)	0.0088 (0.0014)	0.5012 (0.0002)
	AdaBoost	0.7152 (0.0022)	0.2188 (0.0017)	0.5330 (0.0010)
	NB	0.8140 (0.0039)	0.1194 (0.0026)	0.5333 (0.0009)

We observed a significant decrease in performance for SVM, AdaBoost, and Naive Bayes methods in detecting patients. This finding further supports the notion, documented in the literature, that machine learning algorithms may not perform well when dealing with large sample sizes. However, in the case of PRSVM, the AUC increases as the sample size grows. This result indicates that PRSVM has the ability to overcome this issue, showcasing its effectiveness in handling larger datasets.

III. DISCUSSION AND CONCLUSION

The experimental results demonstrate that, in comparison to other machine learning classifiers, PRSVM achieves a lower FNR and higher AUC. These findings indicate that the PRSVM model can more accurately detect patients from imbalanced fMRI datasets. PRSVM demonstrates improved performance as the number of samples from typical controls increases, effectively addressing the challenge of the limited availability of patient data in practical applications. Furthermore, incorporating additional samples from both patients and typical controls into the training set has resulted in significant enhancements, overcoming the limitations of previous machine learning algorithms that were predominantly effective on small datasets. Consequently, we

believe PRSVM holds great value for practical applications. In our future work, we plan to optimize the model's parameters and explore the utilization of kernel methods to achieve even higher classification accuracy. Additionally, we aim to apply this method to a broader range of fMRI data classification tasks.

ACKNOWLEDGMENT

The work by Qiang Wu is partially supported by NSF (DMS-2110826).

REFERENCES

[1] S. Ogawa, T.-M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." proceedings of the National Academy of Sciences, vol. 87, no. 24, pp. 9868–9872, 1990.

[2] C. S. Roy and C. S. Sherrington, "On the regulation of the blood-supply of the brain," The Journal of physiology, vol. 11, no. 1-2, p. 85, 1890

[3] J. Belliveau, D. Kennedy, R. McKinstry, B. Buchbinder, R. Weisskoff, M. Cohen, J. Vevea, T. Brady, and B. Rosen, "Functional mapping of the human visual cortex by magnetic resonance imaging," Science, vol. 254, no. 5032, pp. 716–719, 1991.

[4] B. Sen, G. A. Bernstein, T. Xu, B. A. Mueller, M. W. Schreiner, K. R. Cullen, and K. K. Parhi, "Classification of obsessive-compulsive disorder from resting-state fmri," in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).IEEE, 2016, pp. 3606–3609.

[5] Q. Yu, E. B. Erhardt, J. Sui, Y. Du, H. He, D. Hjelm, M. S. Cetin, S. Rachakonda, R. L. Miller, G. Pearlson et al., "Assessing dynamic brain graphs of time-varying connectivity in fmri data: application to healthy controls and patients with schizophrenia," Neuroimage, vol. 107, pp. 345–355, 2015.

[6] X.-H. Zhao, P.-J. Wang, C.-B. Li, Z.-H. Hu, Q. Xi, W.-Y. Wu, and X.-W. Tang, "Altered default mode network activity in patient with anxiety disorders: an fmri study," European journal of radiology, vol. 63, no.3, pp. 373–378, 2007.

[7] D. D. Cox and R. L. Savoy, "Functional magnetic resonance imaging (fmri)"brain reading": detecting and classifying distributed patterns of fmri activity in human visual cortex," Neuroimage, vol. 19, no. 2, pp. 261–270, 2003.

[8] S. Song, Z. Zhan, Z. Long, J. Zhang, and L. Yao, "Comparative study of svm methods combined with voxel selection for object category classification on fmri data," PloS one, vol. 6, no. 2, p. e17191, 2011.

[9] S. Wang, F. Duan, and M. Zhang, "Convolution-gru based on independent component analysis for fmri analysis with small and imbalanced samples," Applied Sciences, vol. 10, no. 21, p. 7465, 2020.

[10] L. Shao, Y. You, H. Du, and D. Fu, "Classification of adhd with fmri data and multi-objective optimization," Computer Methods and Programs in Biomedicine, vol. 196, p. 105676, 2020.

[11] L. Kanner et al., "Autistic disturbances of affective contact," Nervous child, vol. 2, no. 3, pp. 217–250, 1943.

[12] C. Craddock, Y. Benhajali, C. Chu, F. Chouinard, A. Evans, A. Jakab, B. S. Khundrakpam, J. D. Lewis, Q. Li, M. Milham et al., "The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives," Frontiers in Neuroinformatics, vol. 7, p. 27, 2013.

[13] S. Chen, J. Kang, and G. Wang, "An empirical bayes normalization method for connectivity metrics in resting state fmri," Frontiers in neuroscience, vol. 9, p. 316, 2015.

[14] X. Yang, M. S. Islam, and A. A. Khaled, "Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite abide dataset," in 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2019, pp. 1–4.

[15] X.-a. Bi, Y. Wang, Q. Shu, Q. Sun, and Q. Xu, "Classification of autism spectrum disorder using random support vector machine cluster," Frontiers in genetics, vol. 9, p. 18, 2018.

[16] M. R. Arbabshirani, S. Plis, J. Sui, and V. D. Calhoun, "Single subject" prediction of brain disorders in neuroimaging: Promises and pitfalls," Neuroimage, vol. 145, pp. 137–165, 2017.

[17] X. Yang, P. T. Schrader, and N. Zhang, "A deep neural network study of the abide repository on autism spectrum classification," International Journal of Advanced Computer Science and Applications, vol. 11, no. 4,2020.

[18] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the abide dataset," NeuroImage: Clinical, vol. 17, pp. 16–23, 2018.

[19] L. Shao, C. Fu, Y. You, and D. Fu, "Classification of asd based on fmri data with deep learning," Cognitive Neurodynamics, vol. 15, no. 6, pp. 961–974, 2021.

[20] S. Liu and Q. Wu, "Pairwise learning for imbalanced data classification," in 2021 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2021, pp. 186–189.

[21] Y. Feng, Y. Yang, X. Huang, S. Mehrkanoon, and J. A. Suykens, "Robust support vector machines for classification with nonconvex and smooth losses," Neural computation, vol. 28, no. 6, pp. 1217–1247, 2016.

[22] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney et al., "Advances in functional and structural mr image analysis and implementation as fsl," Neuroimage, vol. 23, pp. S208–S219, 2004.

[23] S. B. Eickhoff, K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, and K. Zilles, "A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data," Neuroimage, vol. 25, no. 4, pp. 1325–1335, 2005