High-accuracy Classification of Attention Deficit Hyperactivity Disorder with $L_{2,1}$-norm Linear Discriminant Analysis

Yibin Tang; Xufei Li; Ying Chen; Yuan Zhong; Aimin Jiang; Xiaofeng Liu

CONTENTS

1 **Backgrounds**

2 **Problems in ADHD Classification**

3 **Classification Framework with $L_{2,1}$-norm LDA**

4 **Experiment Results**

5 **Conclusion**

1. Backgrounds

1 Attention Deficit Hyperactivity Disorder (ADHD)

- **Neurobehavioral disease in children** (5%~7%)
- **Symptomatological diagnosis**

  subjective scoring with Hamilton scales

- **Neurobiological diagnosis**

  objective observations of brain, e.g., EEG, PET, MRI

  biosignal, e.g., ReHo, ALFF,

  Functional Connectivity (FC)

2 ADHD classification

- **Machine learning (ML) with three-phase approach**

➢ **feature selection**

SVM-RFE, LASSO, Elastic net

➢ **feature extraction**

PCA, sparse representation, network features

➢ **classifier design**

SVM, ELM, logistic regression, decision tree

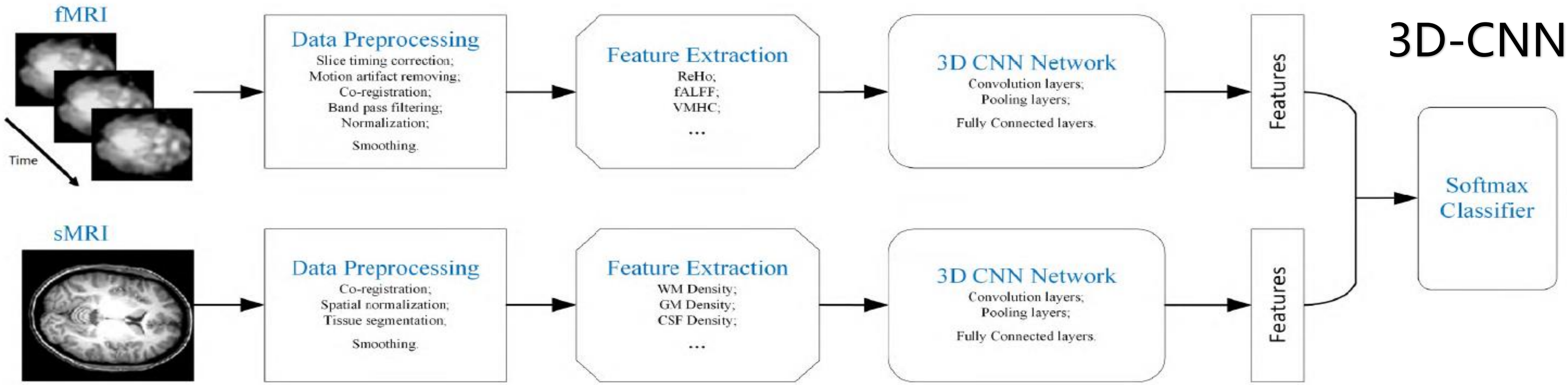
2 ADHD classification

- **Deep learning (DL)**

➢ **remove the barriers among three phases of ML**

➢ **different network to fit given data**

FCNet, Deep fMRI, 3D-CNN, 4D-CNN



3D-CNN

3 Existing ADHD classification performance

- **Accuarcy**
  - 52% ~ 87%
  - Worse performance of DL-based methods
  - **cannot be used in clinical diagnosis**

TABLE 3: Accuracy comparison with various state-of-the-art methods (%).

	NYU	PU	PU_1	KKI	NI
Machine learning					
Fusion fMRI (2017)	52.7	–	85.8	86.7	72.9
L_1BioSVM (2018)	–	81.1	86.7	81.3	–
R-Relief (2019)	70.7	68.6	–	81.8	76.0
Deep learning					
FCNet (2017)	58.5	–	62.7	–	60.0
3D-CNN (2017)	70.5	63.0	–	72.8	–
Deep fMRI (2018)	73.1	–	62.7	–	67.9

2. Problems in ADHD classification

1 Data noise

- **Data collection:** head movement, image aligment error
- **Associate disease:** anxiety disorder, learning disorder

2 Lack of ADHD data

- **Limited subjects (N<250) in ADHD database**
  - **ML:** training data cannot cover the space of test data
  - **DL:** insufficient data to form 'big data'

3. Proposed classification Framework

1 ML framework

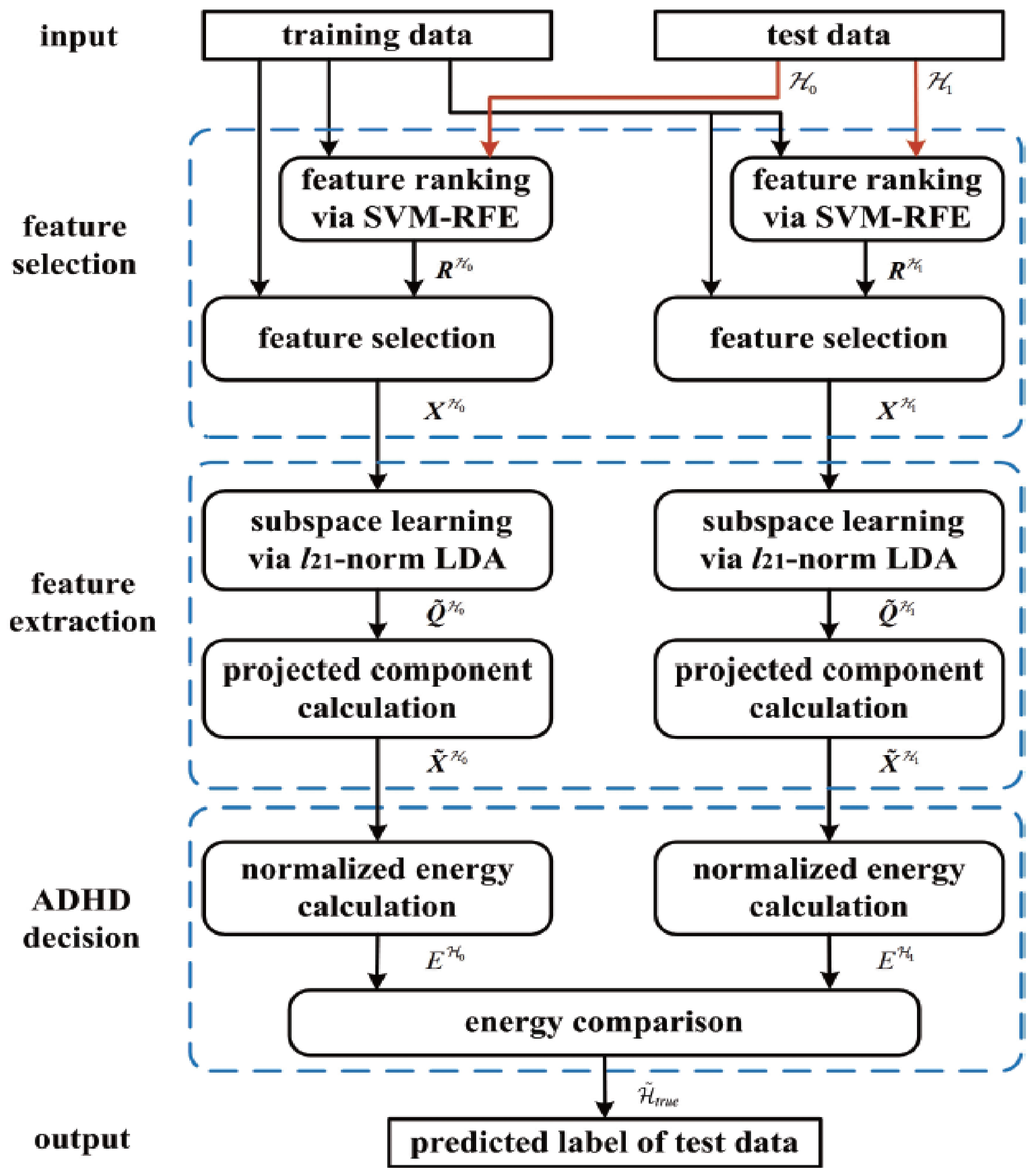
feature selection, feature extraction, ADHD decision

2 Binary hypothesis testing

deal with insufficient data

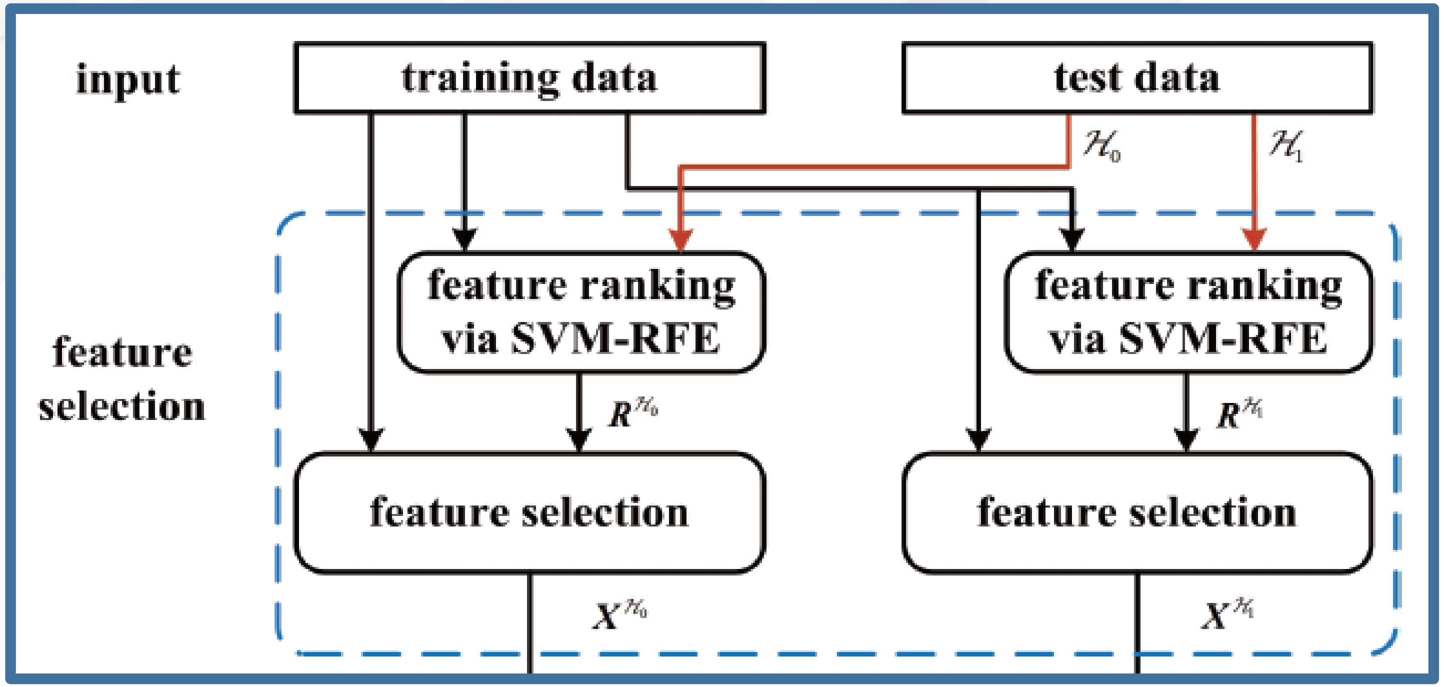
3 $L_{2,1}$-norm LDA

tackle noise disturbance



2 Binary hypothesis

- **FCs of test data affect selected FCs of training data**
  - hypothesized label of test data instead of its true label
  - reliable feature sequence is corresponding to true hypothesis
  - obtain discriminative selected FCs of training data for true hypothesis



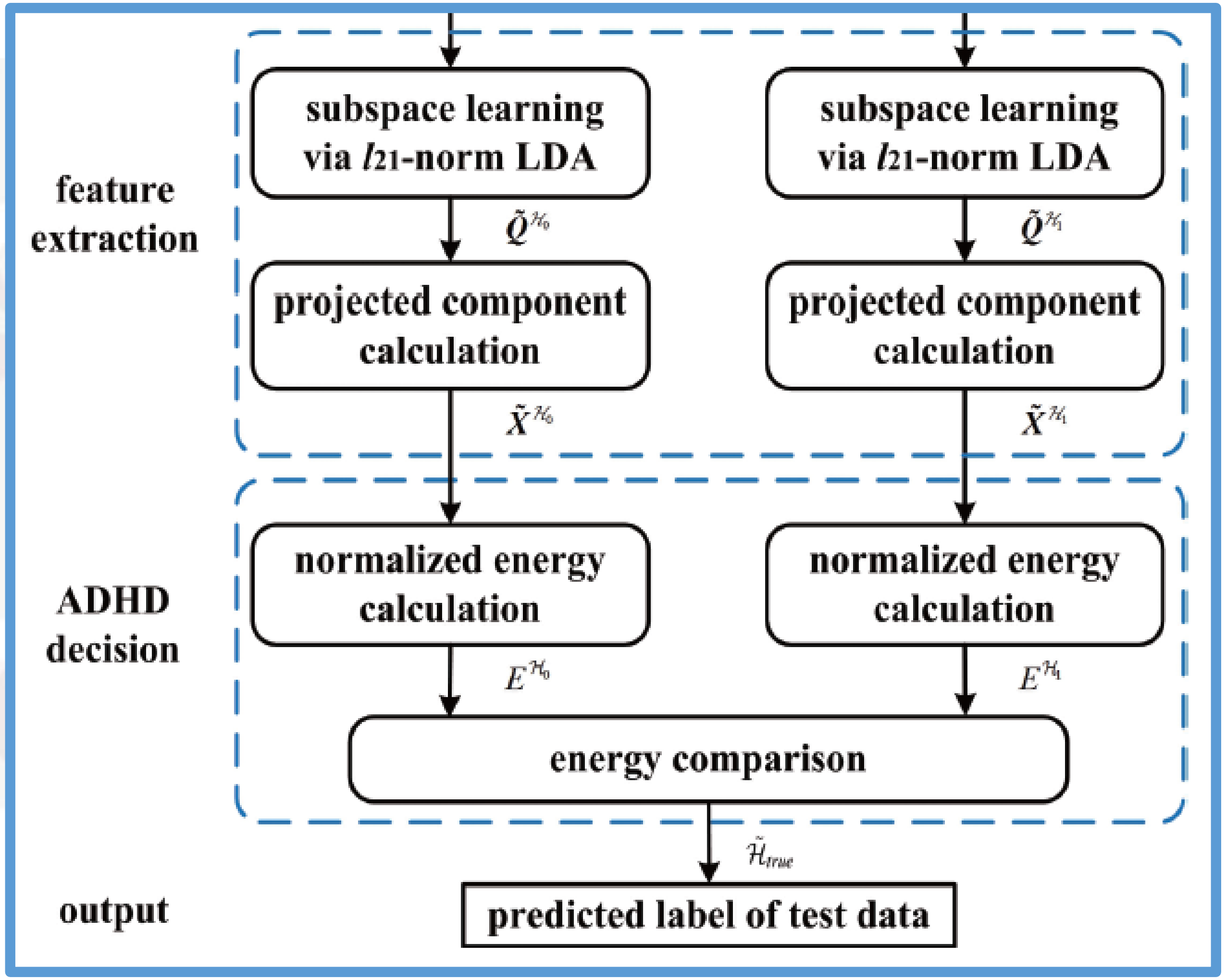
2 $L_{2,1}$-norm LDA

- **Subspace learning to extract ADHD features**

$$\tilde{Q} = \arg \max_{Q^T Q = I} \frac{\|Q^T G_b\|_{2,1}}{\|Q^T G_w\|_{2,1}},$$

$$\begin{cases} \tilde{X}^{\mathcal{H}_0} = (\tilde{Q}^{\mathcal{H}_0})^T X^{\mathcal{H}_0} \\ \tilde{X}^{\mathcal{H}_1} = (\tilde{Q}^{\mathcal{H}_1})^T X^{\mathcal{H}_1} \end{cases}$$

- **ADHD decision**
  - compare subspace energies under binary hypotheses



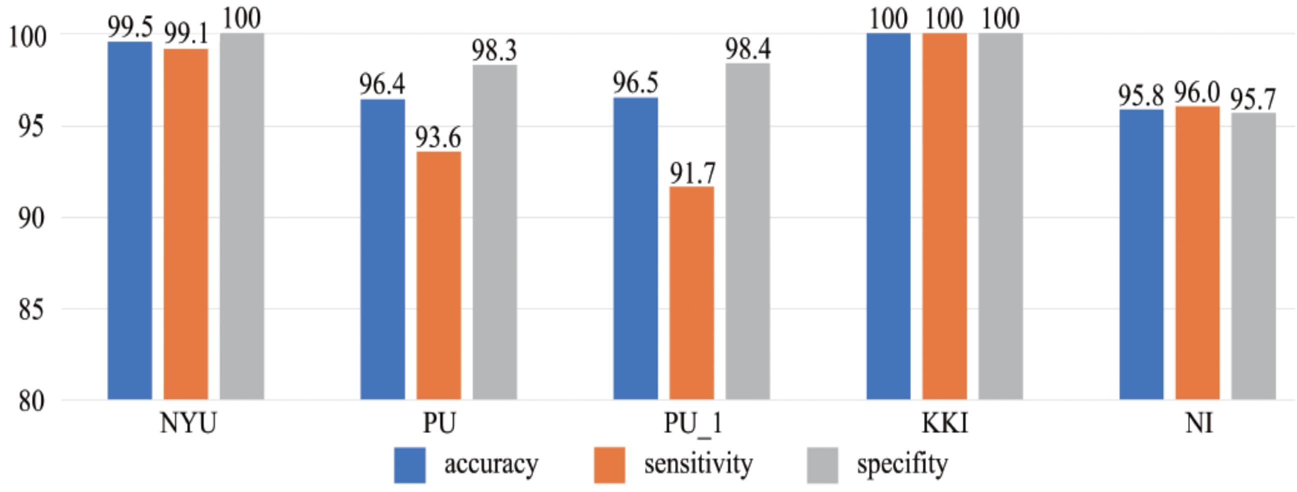
4. Experiment results

1 ADHD accuarcy

Left upper: Databases
Left bottom: Accuracy on various databases
Right: Accurarcy comparison

Site	*Age*	*Female*	*Male*	*Control*	*ADHD*	*Total*
NYU	7-18	77	145	99	123	222
KKI	8-13	37	46	61	22	83
NI	11-22	17	31	23	25	48
PU	8-17	52	142	116	78	194
PU_1*	8-17	36	48	62	24	86

*PU includes three subsets, and PU_1 is the first subset of PU.

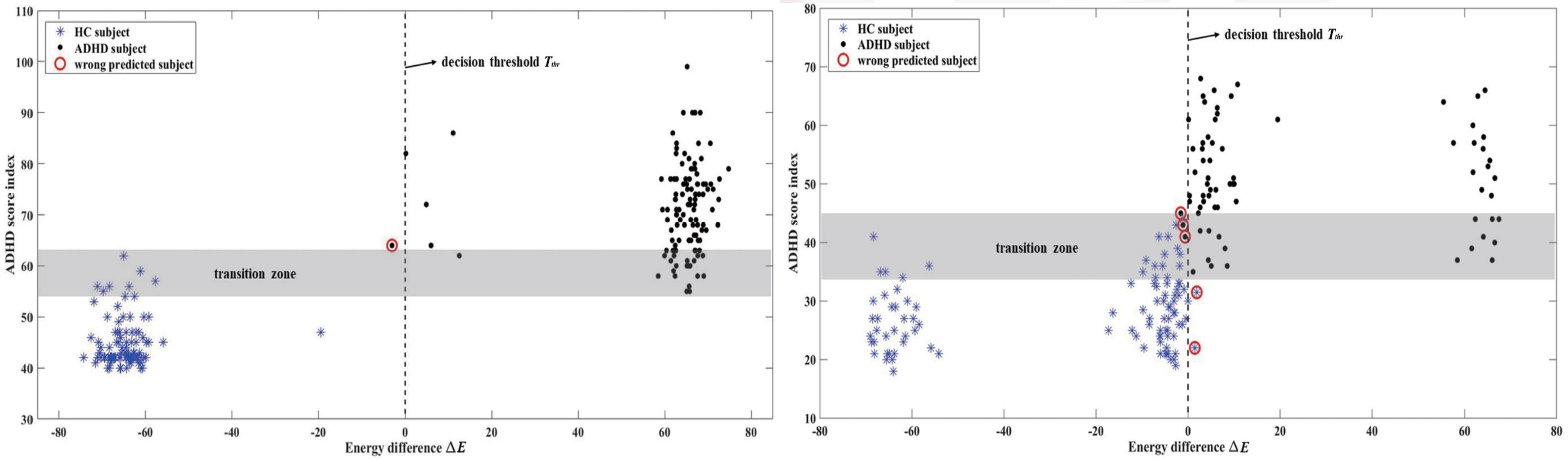


	NYU	PU	PU_1	KKI	NI
Machine learning					
Fusion fMRI (2017)	52.7	–	85.8	86.7	72.9
L_1BioSVM (2018)	–	81.1	86.7	81.3	–
R-Relief (2019)	70.7	68.6	–	81.8	76.0
Deep learning					
FCNet (2017)	58.5	–	62.7	–	60.0
3D-CNN (2017)	70.5	63.0	–	72.8	–
Deep fMRI (2018)	73.1	–	62.7	–	67.9
Binary hypothesis					
SP-BH (2019)	96.2	95.8	91.7	86.7	91.6
Dual-SP-BH (2020)	92.4	92.3	89.4	85.5	81.2
Our method	**99.5**	**96.3**	**96.4**	**100**	**95.8**

4. Experiment results

2 Verify the validity with ADHD symptom score

- **Classification result in line with symptom score**



(a) NYU (b) PU

Evaluation on ADHD decision with ADHD score index.

(We show Fig. 2 of original paper in a clearer form, and give an additional experiment on NYU)

5. Conclusion

1 **Propose a high-accuarcy ADHD classification method**

2 **Classification framework with used appraoches**

- **Binary hypothesis**

  use FCs of test data to affect training data

- **$L_{2,1}$ LDA**

  robust feature learning in subspace

3 **Results**

- **High accuarcy**

  97.6%, better than existing ML and DL methods

- **Prove the validity with ADHD symptom score**

Thank you
for your attention !