

Research Article

A Pilot Study of Diabetes Mellitus Classification from rs-fMRI Data Using Convolutional Neural Networks

Yunfei Liu ¹, Xian Mo,¹ Hao Yang,² Yan Liu ¹ and Junran Zhang ¹

¹Department of Automation, College of Electrical Engineering, Sichuan University, Chengdu, China

²Information Center, West China Hospital, Sichuan University, Chengdu, China

Correspondence should be addressed to Yan Liu; liuyan77@scu.edu.cn and Junran Zhang; junranzhang@scu.edu.cn

Received 23 July 2020; Revised 21 September 2020; Accepted 28 September 2020; Published 21 October 2020

Academic Editor: Jianbing Ma

Copyright © 2020 Yunfei Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. As a chronic progressive disease, diabetes mellitus (DM) has a high incidence worldwide, and it impacts on cognitive and learning abilities in the lifetime even in the early stage, may degenerate memory in middle age, and perhaps increases the risk of Alzheimer's disease. **Method.** In this work, we propose a convolutional neural network (CNN) based classification method to help classify diabetes by distinguishing the brains with abnormal functions from the normal ones on resting-state functional magnetic resonance imaging (rs-fMRI). The proposed classification model is based on the Inception-v4-Residual convolutional neural network architecture. In our workflow, the original rs-fMRI data are first mapped to generate amplitude of low-frequency fluctuation (ALFF) images and then fed into the CNN model to get the classification result to indicate the potential existence of DM. **Result.** We validate our method on a realistic clinical rs-fMRI dataset, and the achieved average accuracy is 89.95% in fivefold cross-validation. Our model achieves a 0.8690 AUC with 77.50% and 77.51% sensitivity and specificity using our local dataset, respectively. **Conclusion.** It has the potential to become a novel clinical preliminary screening tool that provides help for the classification of different categories based on functional brain alteration caused by diabetes, benefiting from its accuracy and robustness, as well as efficiency and patient friendliness.

1. Background

Diabetes affects more than 451 million people (18–99 years) worldwide, and this figure will rise to 693 million by 2045; more surprisingly still, it is estimated that almost half of all people living with diabetes are undiagnosed [1]. It is a group of metabolic diseases characterized by hyperglycemia and frequently accompanied by complications. For many years, the most well-known complications caused by diabetes are dysfunction and failure of organs, like kidney, retina, peripheral nerve, and vasculature [2]. Children diagnosed with T1DM are more likely to perform poorly in school than their nondiabetic classmates and are particularly vulnerable showing impaired results on cognitive tests, learning abilities, and affecting memory [3]. There are now several studies demonstrating a linkage between T2DM and mild cognitive impairment (MCI) and Alzheimer's disease (AD). The co-existence of cerebrovascular disease and T2DM enhances the

correlation with MCI and the development of dementia [4]. Compared to nondiabetic individuals, several studies have demonstrated an increase of AD in T2DM patients [5]. Nowadays, clinical studies have shown that much more attention should be paid to the brain complications of diabetes, like diabetic encephalopathies and cognitive dysfunction. Diabetic encephalopathies are now accepted complications of diabetes, which manifest themselves as a gradual decline of cognitive function and result in brain structural lesions (neural slowing, increased cortical atrophy, microstructural abnormalities in white matter tracts) [6, 7]. There is a growing literature indicating that individuals with diabetes have impairments in recent memory [8], and the mechanism might be due to the fact that glucose transport is significantly reduced in diabetic animals [9]. Most of the current research has formed a consensus that DM affects brain function and brain structure, and the changes in brain function often precede those in brain

structure. Previous studies have shown that white matter could alter after a six-week training period [10]. That is to say, the impact of diabetes on the brain is initiated from brain function and gradually extends to the brain structure. The reason might be that the long-term abnormal blood glucose level damages the cognitive function through its negative effect on target organs, namely, the central nervous system [11], which may help explain why animals that have abnormal glucose metabolism have more hippocampal damage [12]. Glucose borne by the blood accounts for 99% of the brain energy requirements, and metabolic substrate delivery also may influence brain function and structure [13]. What is more serious is that the damage is progressive and irreversible; in most cases, it will develop to dysfunction. It will cause patients to suffer long-term pain and seriously affect the quality of life. A few studies have indicated that reasonably controlling blood glucose with antidiabetic treatments on time may help prevent the dysfunction in diabetic patients [14]. Accordingly, the detection of diabetes mellitus and other changes in brain function can reduce the production probability of those serious complications; it also plays an important role in treatment planning and makes a positive influence on prognosis.

As diabetic encephalopathy is a degenerative disease of the nervous system, we can choose to directly observe the degree of cognitive impairment of patients through modern imaging methods such as magnetic resonance imaging (MRI), so as to explore the correlation between it and diabetes. MRI imaging technology can be divided into structural modalities and functional modalities. Compared with structural images, functional images are more sensitive to early brain changes and so can better reflect earlier changes in brain function. fMRI uses oxygen consumption in brain tissue to determine whether a part of the brain is active in the resting state. The study found that the resting brain was not completely calm but had spontaneous BOLD signal fluctuations which account for 60 to 80 percent of the total energy consumed by the brain. A large number of studies have also shown that the inherent spontaneous neural activity signals of the brain at rest have important physiological significance [15]. Besides, resting-state functional MRI (rs-fMRI), which reflects the variation characteristics of spontaneous resting-level activities in resting state, has been increasingly applied in the field of brain science with unparalleled advantages recently. rs-fMRI is the preferred modality to investigate brain function due to its high temporal resolution and has already been used to measure spontaneous brain activity in patients with diabetes and to reflect changes in brain functional damage caused by diabetes encephalopathy [11]. We can observe the dysfunction degree directly through modern imaging modalities in a noninvasive way and explore its correlation with diabetes and its complication. By using rs-fMRI, researchers have demonstrated that diabetes is related to different indices of functional brain alterations, including regional homogeneity (ReHo) and amplitude of low-frequency fluctuations (ALFF) [16]. ALFF analysis is an important method for depicting the various characteristics of global rs-fMRI signals through measuring the intensity of neural

activity at the single-voxel level and evaluating the differences in the amplitude of low-frequency oscillations of brains [17]. Abundant ALFF studies have demonstrated that the brain function of diabetic patients has changed. Previous studies by Cui et al. [16] showed that patients with T2DM had significantly decreased ALFF values in the postcentral gyrus and occipital lobe. Patients performed worse on several cognitive tests; the researchers speculated that this impaired cognitive performance was correlated with decreased activity in the cuneus and lingual gyrus in the occipital lobe. Recent studies by Xia et al. [18] indicated that T1DM patients showed significantly decreased ALFF values in the posterior cingulate cortex (PCC) and right inferior frontal gyrus compared with the healthy controls. Furthermore, they found a positive correlation between decreased ALFF values in the PCC and Rey-Osterrieth Complex Figure Test- (CFT-) delay scores in T1DM patients.

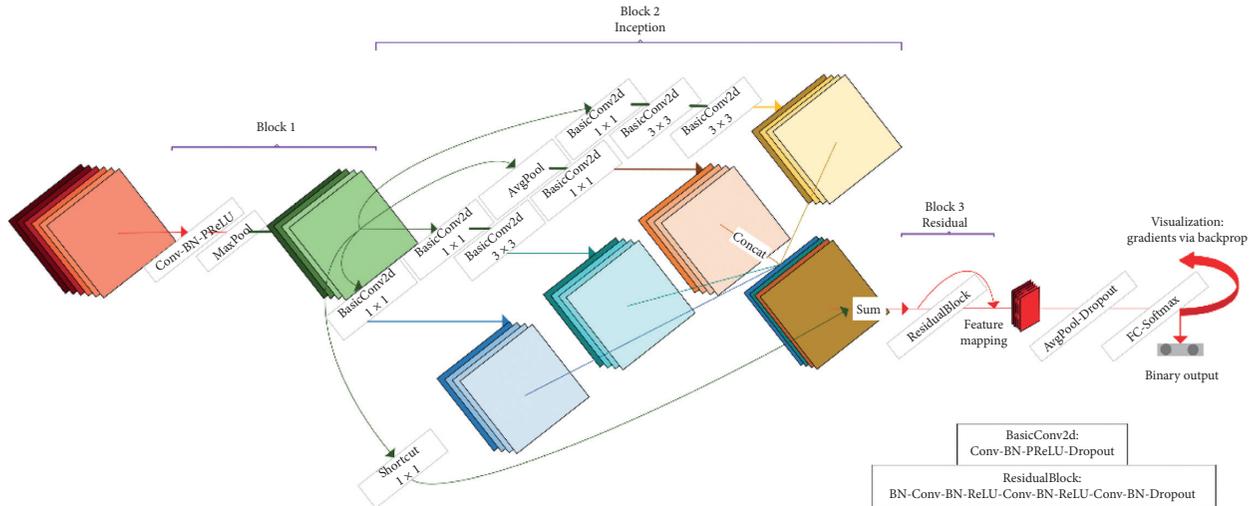
Although ALFF has already been widely used in brain function research, the main analysis still follows the traditional ways (statistical analysis, functional connectivity analysis, and correlation analysis) depending on well-trained experts and is always qualitative and subjective. By contrast, we try to analyze the rs-fMRI sequence through deep learning technology and reflect the quantitative relationship between brain dysfunction and diabetes. In this work, we propose a learning-based classification model to distinguish the abnormal ALFF signals from the normal ones. We employ a convolutional neural network architecture to construct our model. The entire pipeline of the proposed method consists of three successive blocks, as shown in Figure 1(a). This fully automated solution can process thousands of heterogeneous images quickly for accurate, objective diabetes detection. Furthermore, we seek to characterize the association between DM and brain function and structure. We study the differences in the brain function changes caused by diabetes between diabetic patients and healthy control groups to prove the reliability of the classification. All information learned in our end-to-end algorithmic pipeline is visualized through the Gradient-weighted Class Activation Mapping (Grad-CAM), and the subregions within the classified image are highlighted intuitively to further observe the extent of diabetes affecting different brain regions. The differential brain regions reflect the influence of diabetes on brain function and structure and provide some insights for the study of the brain complications of diabetes.

2. Methods

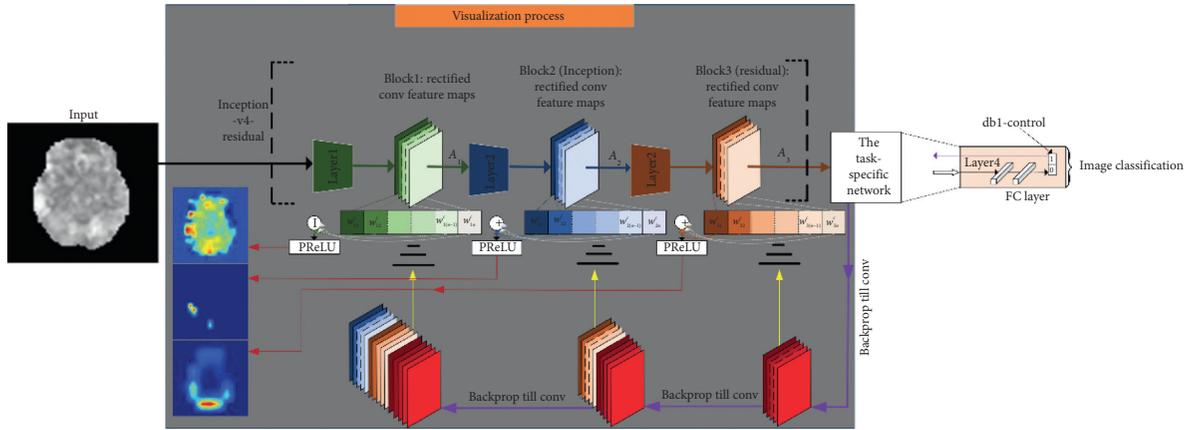
The pipeline of the proposed method is represented in Figure 2.

2.1. Data

2.1.1. Data Acquisition. Our retrospective study includes 47 patients with type 1 diabetes mellitus (denoted as "T1DM"), 73 patients with type 2 diabetes mellitus (denoted as "T2DM"), and 50 healthy controls (denoted as "HC"). Physicians make all the diagnoses according to the criteria



(a)



(b)

FIGURE 1: The abstraction of the proposed algorithmic pipeline and visualization process. (a) The proposed Inception-v4-Residual network with seven parameter layers (about 0.09 million parameters). (b) Grad-CAM overview: Given an image and a class of interest as input, we forward-propagate the ALFF image through the CNN part of the model and then through the task-specific computations to obtain a raw score for the category. Then, the gradients are set to 0 for all classes except the desired class, which is set to 1. Finally, this signal is back-propagated to the parametric rectified convolutional feature maps of interest, which we combine to compute the coarse Grad-CAM localization (blue heatmap), which represents where the model has to look to make the particular decision.

from the American Diabetes Association [19]. All subjects are right-handed and undergo brain scans at the Huaxi MR Research Center of the West China Hospital. Exclusion criteria for all participants include a history of substance or alcohol abuse, a psychiatric or neurological disorder unrelated to diabetes, contraindications to MRI, and a history of a brain lesion such as tumor or stroke. Control subjects are excluded if they have a fasting blood glucose level ≥ 7.0 mmol/L, glucose level ≥ 7.8 mmol/L after oral glucose tolerance test (OGTT).

All rs-fMRI images are acquired via a Siemens (Erlangen, Germany) 3-Tesla Trio scanner. Subjects are instructed to relax, to keep their eyes closed but to remain awake, to keep their heads still during the scanning, and to avoid thinking of

anything in particular. The functional images are recorded using an echo-planar imaging (EPI) sequence with the following parameters: repetition time/echo time (TR/TE) = 2000/30 ms, 90° flip angle, 240×240 mm field of view (FOV), slice thickness/gap = 5/0 mm, $3.75 \times 3.75 \times 5$ mm³ voxel size, and 64×64 matrix resolution. No very obvious structural damage is found in any subject based on MRI scans, which are examined by two experienced specialists.

2.1.2. Functional Image Preprocessing. The raw rs-fMRI data are preprocessed with the Statistical Parametric Mapping software (SPM8) on the MATLAB (R2013b)

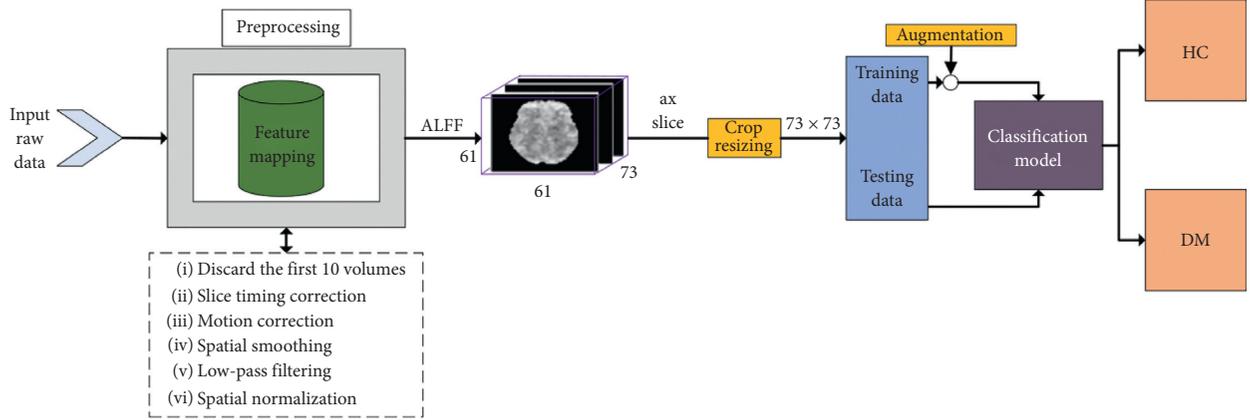


FIGURE 2: The pipeline of the proposed method.

platform. The adopted preprocessing operation is shown in Figure 1. The first 10 volumes of the scanning sessions are discarded due to the instability of the initial MRI signal and adaptation of the participants to the situation. The remaining volumes with size $61 \times 73 \times 61 \times 170$ are analyzed. For each subject, the remaining rs-fMRI images first undergo section-timing correction and are then are realigned. Here, participants with head displacement >2 mm in any x , y , and z direction and $>2^\circ$ in any angular dimension are excluded. After realignment, the resulting images are spatially normalized using 3 mm isotropic voxels. The last step is to smooth the images using a Gaussian kernel of 4 mm full width at half-maximum to improve signal-to-noise ratio and reduce the difference between individuals after standardization.

2.1.3. Feature Mapping. ALFF measures the intensity of neural activity at the single-voxel level and can be calculated through the Data Processing Assistant for Resting-State fMRI (DPARSF) [20] (<http://rfmri.org/DPARSF>). After the above preprocessing, band-pass filtering (0.01–0.08 Hz) is performed on the time series of each voxel to remove the effect of low-frequency drift and high-frequency respiratory and cardiac noise [21]. For each subject, the time series of each voxel is transformed into the frequency domain using Fast Fourier Transform in turn, and the power spectrum is obtained. Then, the square root of the power spectrum is computed and averaged across a predefined frequency interval. The averaged square root is ALFF, an effective indicator of spontaneous neuronal or regional intrinsic activity in the brain [22]. Following the above steps, ALFF, the function map is generated for each subject.

2.1.4. Training and Testing Dataset. Each participant is associated with a diagnostic label of 1 or 0 referring to DM or HC, confirmed by medical specialists. Our data augmentation protocol is as follows. Firstly, all images are normalized to $[0, 1]$ and resized to the standard resolution of 73×73 . Secondly, images are flipped horizontally or

vertically to capture the reflection invariance. The final data transformation is brightness adjustment with one random scale factor α per image, sampled from a uniform distribution over $[-0.2, 0.2]$. We add above transformations that extend the translation invariance, improve our model's ability to generalize, and correctly classify images without a loss of accuracy. To address the imbalance of dataset categories, about 20% of the subjects from HC and DM are assigned to the testing data, and the remaining are used for training.

2.1.5. Model and Visualization. A full diagram of the classification model can be viewed in Figure 1(a), and the abstraction of the visualization process is represented in Figure 1(b).

2.1.6. Classification Model. The network used in this study is inspired by the CNN architecture, Inception-v4-Residual, presented in [23]. It speeds up the flow of information, extracts features from different scales, accelerates network training, and avoids gradient disappearing. Limited by the amount of data, we simplify the original Inception-v4-Residual architecture into three integrated convolutional blocks (block1, block2 (Inception), block3 (Residual)), one average pooling (AvgPool) layer, and one fully connected (FC) layer, as depicted in Figure 1(a). All integrated blocks are stacked by basic blocks as follows.

$$\text{Block1} = \text{Conv} - \text{BN} - \text{PReLU} - \text{MaxP}, \quad (1)$$

$$\left\{ \begin{array}{l} \text{BasicConv} = \text{Conv} - \text{BN} - \text{PReLU} - \text{Dropout} \\ \text{branch1} = \text{BasicConv} \\ \text{branch2} = \text{BasicConv} - \text{BasicConv} \\ \text{branch3} = \text{BasicConv} - \text{BasicConv} - \text{BasicConv} \\ \text{branch4} = \text{AvgP} - \text{BasicConv} \end{array} \right. ,$$

$$\text{Block2} = \text{Concat}(\text{branch1}, \text{branch2}, \text{branch3}, \text{branch4}), \quad (2)$$

$$\begin{aligned} \text{Block3} = & \text{BN} - \text{Conv} - \text{BN} - \text{ReLU} - \text{Conv} - \text{BN} - \text{ReLU} \\ & - \text{Conv} - \text{BN} - \text{Dropout}, \end{aligned} \quad (3)$$

where (branch1, branch2, branch3, branch4) in (2) refers to the concatenation of the feature maps produced in branch 1, \dots , 4. *Conv* is the convolutional layer which learns iteratively filters that transform the input into hierarchical feature maps. *BN* is the normalization layer used to normalize the activation for fast convergence and to improve performance [24]. A BN layer conducts an affine transformation with the following equation:

$$y = \gamma x + \beta, \quad (4)$$

where γ and β are learned for every activation in feature maps. Parametric Rectified Linear Unit (PReLU) [25] or Rectified Linear Unit (ReLU) [26] layer applies an elementwise activation function, such as the max ($\alpha x, x$) or max ($0, x$) thresholding, and does not change the size of the image volume, as defined by

$$x_l = (\text{PReLU}(\text{conv}_l(x_{l-1}))), \quad (5)$$

where conv_l represents a convolutional layer l , which returns the previous convolutional layer's output volume. The pooling layer (MaxP or AvgP) performs a downsampling operation along the spatial dimensions to reduce hyperparameters and prevent overfitting. The key idea of dropout layer (dropout) [27] is to randomly drop units (along with their connections) from the neural network during training. This strategy prevents too much coadaptation of units, significantly reduces overfitting, and gives major improvements over other regularization methods. FC layer computes the class scores, resulting in a volume of several classes. As the name implies, each neuron in the FC layer will be connected to all the numbers in the previous volume [28, 29]. An abstraction of this feature-learning architecture is represented in Figure 1(a), and see Table 1 for detailed architectures.

2.1.7. Implementation. Our implementation for Inception-v4-Residual follows the practice in [30, 31]. We initialize the weight as in [25] and train the Inception-v4-Residual network from scratch. We use Adam and SGD in the early and late stages of model training with a minibatch size of 10, respectively. The learning rate (lr) starts from 0.001, and the model is trained for up to 20×10^4 iterations. Figure 3 shows how the learning rate changes and how the optimizer is selected during training. We use weight decay of 0.0005 and momentum of 0.8. Since the data are relatively unbalanced, we also use the 2-class categorical focal loss for discrimination [32]. In practice, we use an α -balanced variant of the focal loss:

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t), \quad (6)$$

where $t \in \{0, 1\}$ specifies the ground-truth class and $p_t \in [0, 1]$ is the model's estimated probability for the class with a label t . A weighting factor $\alpha \in [0, 1]$ is applied for class

$t = 0$ and $1 - \alpha$ for class $t = 1$. In practice, α is set by inverse class frequency, and we set $\gamma = 2$ in our experiment.

2.1.8. Visualization Process. The efficient visualization process used in this work is Grad-CAM proposed by Selvaraju et al. [33]—for making any CNN based models more transparent by producing “visual explanations.” Grad-CAM uses the gradients of any target concept, flowing into the corresponding convolutional layer to produce a coarse localization map highlighting the most important regions in the image for predicting the concept. We apply Grad-CAM, a class-discriminative localization technique, to find characteristic brain regions that can classify diabetic patients. The procedure for generating these maps is illustrated in Figure 1(b).

As shown in Figure 1(b), to obtain the class-discriminative localization map Grad-CAM $L_{n, \text{Grad-CAM}}^c \in \mathbb{R}^{\mu \times \nu}$ of width $\mu = 73$ and height $\nu = 73$ for any class $c = 2$, we first compute the gradient of the score for the class c , y^c with respect to feature maps A_n^k (A_1^k, A_2^k, A_3^k) of a convolutional block, i.e., $\partial y^c / \partial A_n^k$. These gradients flowing back are global-average-pooled to obtain the neuron importance weights $\alpha_{n,k}^c$. The final localization map is calculated through (7) and visualized as a heatmap in Figure 1(b). The highlighted regions in the ALFF image might be used to help real-time clinical validation of automated detection in the future.

$$L_{n, \text{Grad-CAM}}^c = \text{PReLU} \left(\underbrace{\left(\underbrace{\left(\underbrace{\frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{n,i,j}^k}}_{\text{backprop gradients}} \right)}_{\text{global average pooling of } \alpha_{n,k}^c} \right)}_{\text{linear combination}} \cdot A_n^k \right). \quad (7)$$

3. Results

We use 5-fold cross-validation in this study. Average metrics originated from 5 test runs on respective held-out data.

3.1. Local Cross-Validation Results. In our work, we mainly use the accuracy for evaluating our network architecture. All the networks are trained using the same experimental setup as in Section 2.1.7. The initial lr is set to 0.001, and we train models for 300 epochs with a batch size of 10 with the same data augmentation strategies as in Section 2.1.4. Note that we use early stopping [34] to prevent overfitting in all the networks. The accuracy of the single-crop is provided in Table 2 for evaluating our models. Our algorithm scores an AUC of 0.8690 during cross-validation and also achieves an average 77.50% sensitivity and a 77.51% specificity. This ROC curve is plotted in Figure 4. In the training phase, accuracy and also loss of training and testing are measured as is depicted in Figure 5. The experimental results in Figure 5 show that the Inception-v4-Residual network with only 0.091 M trainable parameters converges fairly fast, which indicates that our network has a real-time runtime performance and has computational efficiency. We empirically

TABLE 1: Inception-v4-Residual architectures. Building blocks are shown in brackets (see also Figure 1(a)).

Layer name	Output size	Conv
Block1	37×37	$[3 \times 3, 16, \text{stride } 2, \text{padding } 1]$
	19×19	$[3 \times 3, \text{max pool, stride } 2]$
	19×19	$[1 \times 1, 16, \text{stride } 1]$
Block2 inception	19×19	$\left[\begin{array}{l} 1 \times 1, 16, \text{stride } 1 \\ 3 \times 3, 32, \text{stride } 1, \text{padding } 1 \end{array} \right]$
	19×19	$\left[\begin{array}{l} 1 \times 1, 16, \text{stride } 1 \\ 3 \times 3, 16, \text{stride } 1, \text{padding } 1 \\ 3 \times 3, 32, \text{stride } 1, \text{padding } 1 \end{array} \right]$
	19×19	$\left[\begin{array}{l} 3 \times 3, \text{stride } 1, \text{padding } 1 \\ 1 \times 1, 8, \text{stride } 1 \end{array} \right]$
cat	branch_all = Concat[branch1, branch2, branch3, branch4] = 88	
sum	$\left[\begin{array}{l} \text{shortcut} = 1 \times 1, 88, \text{stride } 1 \\ \text{shortcut} + \text{branch_all} \end{array} \right]$	
Block3 residual	9×9	$\left[\begin{array}{l} 1 \times 1, 44, \text{stride } 1 \\ 3 \times 3, 44, \text{stride } 2 \\ 1 \times 1, 88, \text{stride } 1 \end{array} \right] + [3 \times 3, 88, \text{stride } 2]$
# Params.	Adaptive average pool, 2d FC, softmax 0.09 M	
FLOPs	1×10^7	
Trunk depth	7	

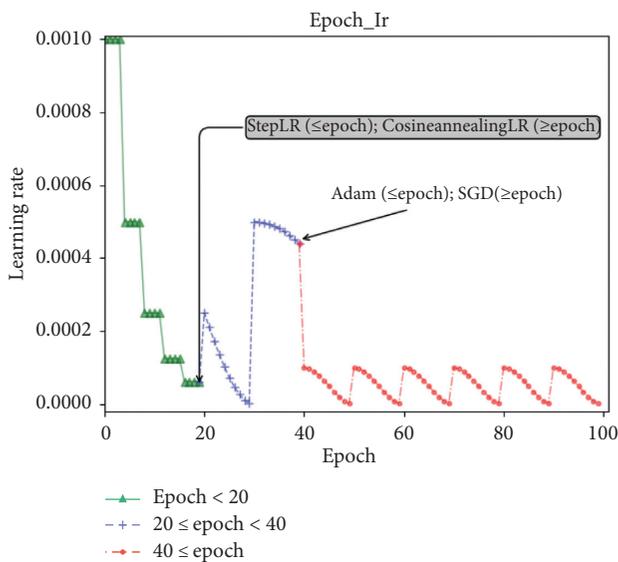


FIGURE 3: Schematic diagram of optimizer selection and dynamic adjustment of learning rate. Green: the initial lr decays by $\gamma = 0.5$ every four epochs. Blue and red: after the 20th epoch, cyclical lr is a function of epoch. Epoch ≤ 40 : Adam. Epoch > 40 : SGD.

demonstrate the Inception-v4-Residual network's effectiveness on ALFF classification accuracy grouped by subject in Figure 6. Through Figure 6, we can see that the recognition rate of the model to diabetic patients is quite high. However, classification for HC is slightly worse because of the lack of distinct image features. Compared with false-negatives, false-positives could be expected when using an automated method for the detection of a patient.

3.2. Grad-CAM Visualization Results. For efficiently triaging referrals and focusing on one's clinical examination, it is highly important to interpret the output of detection-guiding software reasonably. Toward that end, we apply the Grad-CAM visualization method to locate the most discriminative features of our deep learning network. Figure 7 illustrate some feature maps of the final filter (kernel) of block1, block2, and block3, where the regions that contributed most to the final classification result are highlighted on the heatmaps. Those highlighted regions tie the mathematical learning of the network to the domain of clinical data. The distribution differences of these important regions corroborate the domain-guided learning procedure of our model. The average

TABLE 2: Accuracies achieved from Inception-v4-Residual network across five runs.

Inception-v2-Residual	Training time per epoch (s)	Total time (m)	Top-1 accuracy (%)	AUC
run1	5.65	9.51	89.51	0.8708
run2	5.64	9.49	89.88	0.8862
run3	5.71	9.61	90.12	0.8721
run4	5.70	5.60	90.12	0.8762
run5	5.59	12.03	90.12	0.8391
# of params.		=0.09 M		
Mean accuracy		89.95%		
Mean AUC		0.8690		

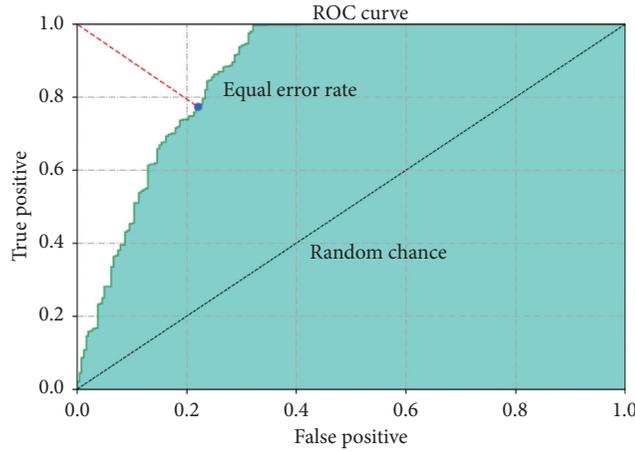


FIGURE 4: The average ROC curve derived from the Inception-v4-Residual network’s run3. The black dotted line represents the trade-off resulting from random chance. The red curve represents the model’s trade-off, with the blue dot marking the threshold point.

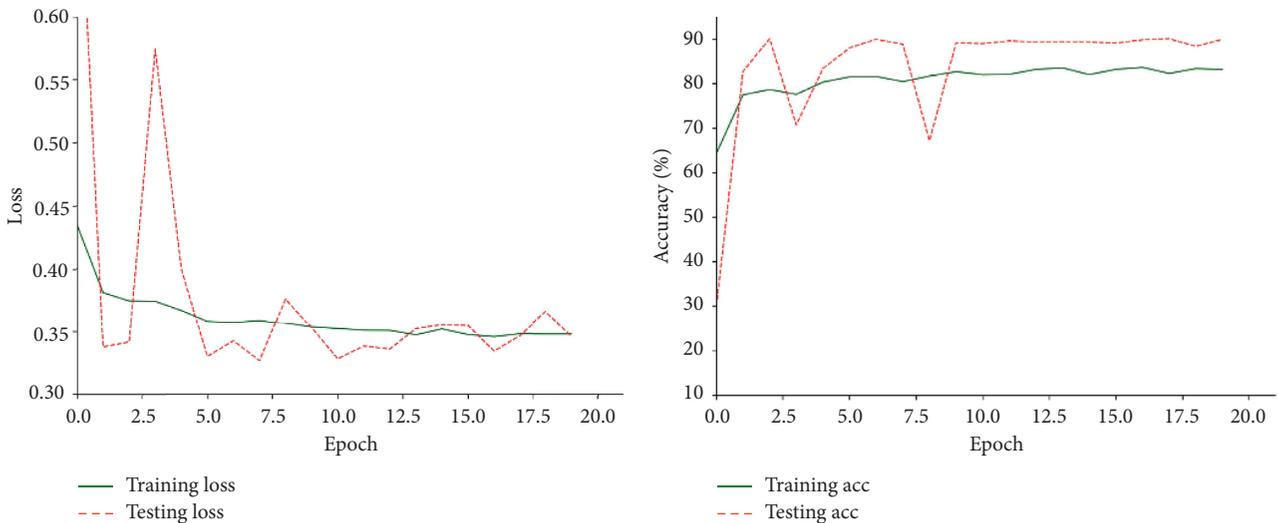


FIGURE 5: Loss and accuracy of training and testing for Inception-v4-Residual network’s run3.

activation of HC is much higher than DM, which proves to some extent the reliability of the classification results.

3.3. Effectiveness of Different Network Architectures. Apart from the Inception-v4-Residual network, we explore several other CNN models, including ResNet [35] and

Inception [36] for performance comparison. The total number of parameters in these networks is shown in Table 3. The same image preprocessing is adopted for a fair comparison. We train network entirely for several epochs at a certain resolution from scratch. Table 3 shows the experiment results. It is observed that the network architectures have a slight impact on the performance, and Inception-v4-Residual

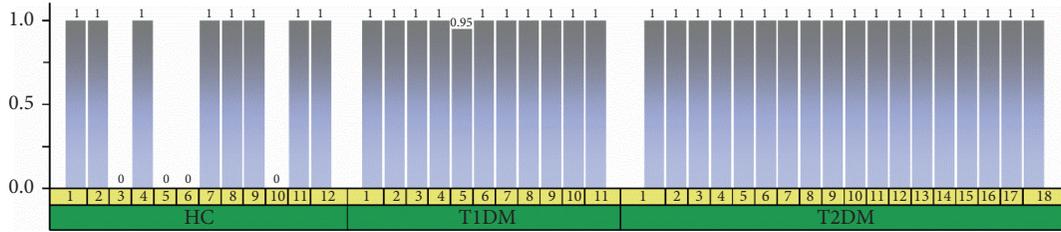


FIGURE 6: Subject-level ALFF classification accuracy on the test set using the best model.

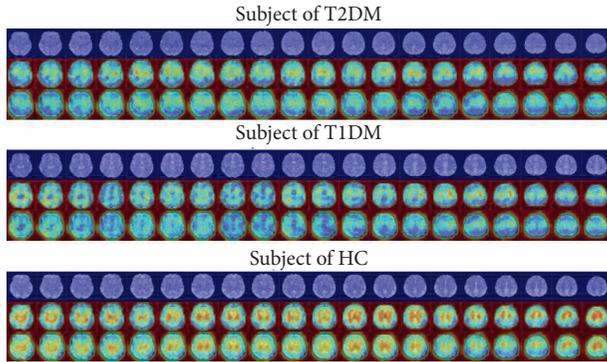


FIGURE 7: Examples of visualization map generated from deep features. Note that, in all subjects, red regions correspond to a high score for the class. The figure is best viewed in color.

TABLE 3: Accuracies achieved from Inception-v4-Residual, ResNet, and Inception models.

Ours	Inception	ResNet	#of params. (M)	Top-1 accuracy (%)
✓			0.091	90.12
	✓		0.152	89.88
		✓	2.039	88.78

network achieves better performance than the other two networks, which demonstrates the efficiency of our method.

3.4. Ablation Studies. The main design choices we make for the training and testing procedure are the data augmentation protocols and test-time rotations. To show their impact, we perform two ablation studies.

3.4.1. Rotation. The experimental setting of the training process remains the same as the above experiment, but during testing the images are only presented in their original orientation or show the direction after being rotated at a certain angle. To improve accuracy, we average the classification scores.

3.4.2. Data Augmentation Protocols. During training, we have not adopted any data augmentation protocols. Meanwhile, we separately test the impact of each of the data augmentation protocols and their combination on the experimental results. The testing procedure is unchanged.

Figure 8 shows results on the preliminary test set for our proposed setting and the two ablation studies. The influence of data augmentation protocols is significant but smaller; we note that the brightness adjustment is beneficial to the model’s

generalization ability, but other data augmentation protocols have reduced the model’s ability to learn features. Not using test-time rotation slightly decreases the mean accuracy (0.24%) and AUC (0.006) in fivefold cross-validation.

4. Discussion

In this work, the convolution neural network based on the classification model is proposed to help DM and brain related abnormalities classification based on rs-fMRI, and the results indicate that it is effective and accurate. In addition to facilitating detection, for the first time, our algorithm pipeline visualizes the abnormal brain areas, which might provide critical information to understand the effect of diabetes on brain dysfunction. Our approach provides an attempt to utilize deep learning to detect DM and brain related abnormalities from resting-state fMRI data. Results obtained suggest that rs-fMRI holds the potential to increase the translation of rs-fMRI data into clinical detection.

In classifying HC and DM, we can achieve an AUC of 0.869 on the local dataset. The visualization results in Figure 7 provide some reliable evidence for the results of this experiment. As we have seen, the degree of activation of HC and DM is significantly different, which confirms the previous research finding that DM patients have a decreased spontaneous brain activity on rs-fMRI. However, the difference in activation between T1DM and T2DM is not that obvious. To a certain degree, this visualization result can explain the high level of accuracy in the experiment. As expected, decreased neural activity is significantly associated with DM, a result which is in agreement with other studies [16]. It is extremely difficult to accurately classify diabetes and its type in the early stages even for medical care personnel because many diabetic individuals do not easily fit into a single class, and assigning a type of diabetes to an individual often depends on the circumstances present at the time of diagnosis. The detection of subtle differences in brain function between T1DM and T2DM poses an important limitation on accurate identification in future DM typing detection systems. In future research, we look forward to combining manual features for targeting specific characteristics of DM and the robust potential of deep learning systems to characterize the type of diabetes accurately to yield more clinically useful results.

Further optimization of the sensitivity metric might be necessary to ensure a minimum false-negative rate for proper clinical application of our algorithm. The computer-aided system for DM classification must minimize false-

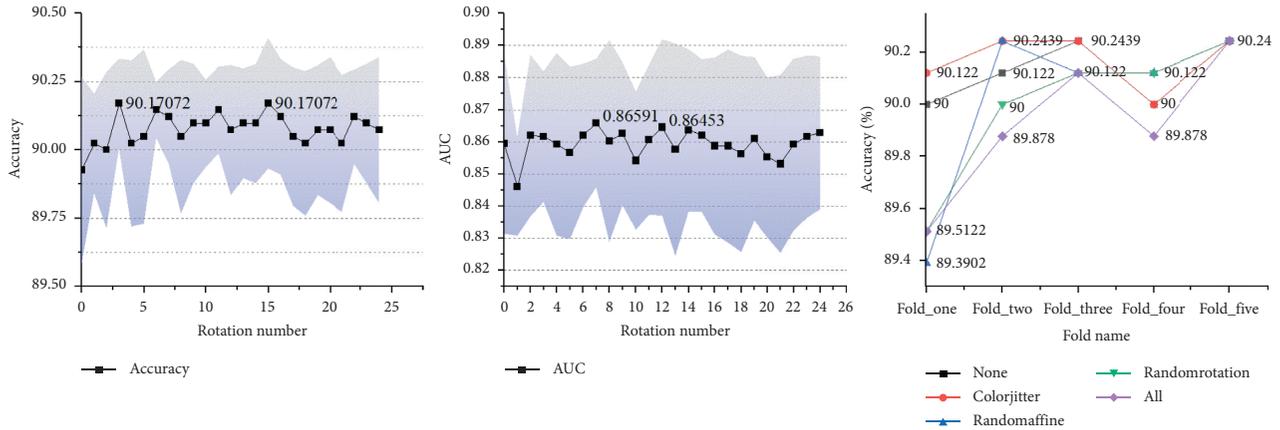


FIGURE 8: Metrics on the preliminary test set measured after the early stopping for the proposed setting and the two ablation studies. The shaded area marks one standard deviation. The figure is best viewed electronically and zoomed in.

negative results to provide necessary glucose care for patients. During clinical use, it may be quite critical to control specific variances in local dataset, such as age, to optimize the model for certain demographics. Patient history, duration of diabetes, symptom type, hemoglobin A1C value, genetic factors, and other clinical data may play a crucial role in investigating different types of common patient metadata that may assist healthcare professionals in making a correct diagnosis of the type of DM. Adding confirmed clinical information into the classification system may yield insightful correlations into underlying DM risk factors outside of imaging information, potentially enhancing classification accuracy between T1DM and T2DM. Several limitations of this study should be noted. First, during subsequent experiments, a large patient and control subject cohort is needed to create a more robust model, and independent testing on external datasets is required to confirm its predictive properties. Second, the impact of different background ethnicity and geographic location of demographics on the classification model needs to be considered. Third, our experiments assume that diabetic patients with chronic insulin deficiency or hyperglycemia may cause corresponding changes in brain function and microstructure. These changes can be reflected in rs-fMRI images, which can use deep learning methods to identify the disease but cannot exclude the presence of diabetes without causing changes in brain function and microstructure, for instance, when diabetes is quite early. In short, much effort is still required to achieve the clinical implementation of a texture-based decision-support system in further research.

Diabetic encephalopathies are now accepted complications of diabetes. They appear to differ in T1DM and T2DM as to the underlying mechanisms and the nature of resulting cognitive deficits [3]. Both types of diabetes are associated with increased risks for micro- and macrovascular disease and cerebrovascular accidents with compounding effects on cognitive deficits [13]. Studies of brain function and structural neuroimaging have demonstrated associated anomalies. As shown in Figure 7, it should be mentioned

though that the class activation mapping in T1DM appears to be different from that of T2DM, which coincides with previous studies. There is evidence suggesting that the progressive deficits in brain function and structure may develop already in patients with prediabetes [13] (T1DM leads to neuronal loss and disintegration of neuronal networking fundamental to cognitive function; T2DM results in neuronal loss). However, some of the underlying pathogenetic mechanisms are different in the encephalopathies of the two types of diabetes. Therefore, continued investigations are needed to start to formulate precise therapeutic interventions to curtail the increase in these major complications. In future research, we will seek to characterize the association between DM and brain function and structure using CNN and Grad-CAM to further quantitatively investigate the impact of diabetes on brain complications. Definitely, this can reduce the production probability of those serious complications; it also plays an important role in treatment planning and makes a positive influence on prognosis.

In this study, we distinguish diabetic patients from healthy controls and seek to characterize the association between DM and brain function and structure. We propose that, among DM patients, the activation degree of the feature map would be associated with decreased spontaneous brain activity. This study may also help create the view that differences in brain activity of different diabetes types are closely related to the corresponding brain complications. On the whole, we propose a cost-effective and time-efficient automatic diagnosis algorithm of diabetes which shows the potential of automated feature-learning systems in streamlining current diabetes screening programs.

5. Conclusion

In this work, we explore an approach based on deep learning to distinguish the DM data from normal control data with 89.95% accuracy. The results showed that rs-fMRI holds great promise for the prediction of DM. However, further validation on independent datasets is required to confirm its

predictive properties. This deep learning solution and our algorithm pipeline provide a new idea for the diagnosis of DM. It will potentially alleviate the workload of manual analysis and guide high-risk patients for referral to further care. We believe that the methodology presented in this work can be also generalized to predict the different types of DM for different age groups, and it can eventually lead to improvements in treatment personalization and patient survival.

Abbreviations

ALFF:	Amplitude of low-frequency fluctuation
HC:	Healthy controls
DM:	Diabetes mellitus
T1DM:	Type 1 diabetes mellitus
T2DM:	Type 2 diabetes mellitus
MRI:	Magnetic resonance imaging
Grad-CAM:	Gradient-weighted Class Activation Mapping
AUC:	Area under the receiver operating characteristic curve
GDM:	Gestational diabetes mellitus
DR:	Diabetic retinopathy
ReHo:	Regional homogeneity
lr:	Learning rate.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors' Contributions

YFL, YL, HY, and JRZ substantially contributed to method and design. XM and YFL participated in data acquisition. YFL was responsible for code and network structure design. YFL, YL, and XM drafted the article. YL and JRZ critically revised the article for important intellectual content. All authors approved the final version to be published.

Acknowledgments

The publication of this article was sponsored by the National Science Foundation of China under Grant 61902264 and the Key Research and Development Projects in Sichuan Province under Grant 2019YFS0125, Sichuan University-Zigong City Cooperation Project 2018CDZG-19.

References

- [1] N. H. Cho, J. E. Shaw, S. Karuranga et al., "IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045," *Diabetes Research & Clinical Practice*, vol. 138, pp. 271–281, 2018.
- [2] A. D. Association, "Classification and diagnosis of diabetes," *Diabetes Care*, vol. 40, no. Supplement 1, pp. S11–S24, 2017.
- [3] A. A. Sima, "Encephalopathies," *The Emerging Diabetic Complications*, vol. 47, no. 4, pp. 279–293, 2010.
- [4] R. Peila and L. J. Launer, "Type 2 diabetes, apoe gene, and the risk for dementia and related pathologies," *The Honolulu-Asia Aging Study*, vol. 51, no. 4, pp. 1256–1262, 2011.
- [5] J. A. Luchsinger, M. Tang, Y. Stern, S. Shea, and R. Mayeux, "Diabetes mellitus and risk of alzheimer's disease and dementia with stroke in a multiethnic cohort," *American Journal of Epidemiology*, vol. 154, no. 7, pp. 635–641, 2001.
- [6] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., "The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [7] R. J. McCrimmon, C. M. Ryan, and B. M. Frier, "Diabetes and cognitive dysfunction," *The Lancet*, vol. 379, no. 9833, pp. 2291–2299, 2012.
- [8] S. Kalmijn, E. J. M. Feskens, L. J. Launer, T. Stijnen, and D. Kromhout, "Glucose intolerance, hyperinsulinaemia and cognitive function in a general population of elderly men," *Diabetologia*, vol. 38, no. 9, pp. 1096–1102, 1995.
- [9] W. M. Pardridge, D. Triguero, and C. R. Farrell, "Down-regulation of blood-brain barrier glucose transporter in experimental diabetes," *American Diabetes Association*, vol. 39, no. 9, pp. 1040–1044, 2013.
- [10] J. Scholz, M. C. Klein, T. E. Behrens, and H. Johansen-Berg, "Training induces changes in white-matter architecture," *Nature Neuroscience*, vol. 12, no. 11, pp. 1370–1371, 2009.
- [11] Z.-G. Li, W. Zhang, and A. A. Sima, "The role of impaired insulin/IGF action in primary diabetic encephalopathy," *Brain Research*, vol. 1037, no. 1–2, pp. 12–24, 2005.
- [12] A. M. Magariños and B. S. McEwen, "Experimental diabetes in rats causes hippocampal dendritic and synaptic reorganization and increased glucocorticoid reactivity to stress," *Proceedings of the National Academy of Sciences*, vol. 97, no. 20, p. 11056, 2000.
- [13] A. Convit, O. T. Wolf, C. Tarshish, and M. J. De Leon, "Reduced glucose tolerance is associated with poor memory performance and hippocampal atrophy among normal elderly," *Proceedings of the National Academy of Sciences*, vol. 100, no. 4, p. 2019, 2003.
- [14] C. Messier, "Impact of impaired glucose tolerance and type 2 diabetes on cognitive aging," *Neurobiology of Aging*, vol. 26, no. 1, pp. 26–30, 2005.
- [15] B. B. Biswal, M. Mennes, X.-N. Zuo et al., "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, p. 4734, 2010.
- [16] Y. Cui, Y. Jiao, Y.-C. Chen et al., "Altered spontaneous brain activity in type 2 diabetes: a resting-state functional MRI study," *Diabetes*, vol. 63, no. 2, pp. 749–760, 2014.
- [17] D. S. Margulies, J. Böttger, X. Long et al., "Resting developments: a review of fMRI post-processing methodologies for spontaneous brain activity," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, no. 5–6, pp. 289–307, 2010.
- [18] W. Xia, Y. C. Chen, Y. Luo et al., "Decreased spontaneous brain activity and functional connectivity in type 1 diabetic patients without microvascular complications," *Cellular Physiology and Biochemistry*, vol. 51, 2018.
- [19] V. Gulshan, L. Peng, M. Coram et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *The Journal*

- of the American Medical Association, vol. 316, no. 22, pp. 2402–2410, 2016.
- [20] C. Yan and Y. Zang, “DPARSF: a MATLAB toolbox for “pipeline” data analysis of resting-state fMRI,” *Frontiers in Systems Neuroscience*, vol. 4, p. 13, 2010.
- [21] M. Lowe, B. Mock, and J. Sorenson, “Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations,” *Neuroimage*, vol. 7, no. 2, pp. 119–132, 1998.
- [22] Z. Yu-Feng, H. Yong, Z. Chao-Zhe et al., “Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI,” *Brain and Development*, vol. 29, no. 2, pp. 83–91, 2007.
- [23] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Proceedings of the Thirty-first AAAI conference on artificial intelligence*, San Francisco, CA, USA, February 2017.
- [24] S. Ioffe and C. Szegedy, “Batch normalization: accelerating deep network training by reducing internal covariate shift,” 2015, <https://arxiv.org/abs/1502.03167>.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 2015.
- [26] V. Nair and G. E. Hinton, “Rectified linear units improve restricted Boltzmann machines,” in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, Haifa, Israel, June 2010.
- [27] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [28] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [29] I. Arel, D. C. Rose, and T. P. Karnowski, “Deep machine learning—a new frontier in artificial intelligence research [research frontier],” *IEEE Computational Intelligence Magazine*, vol. 5, no. 4, pp. 13–18, 2010.
- [30] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, <https://arxiv.org/abs/1409.1556>.
- [31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proceedings of the NIPS Proceedings Advances In Neural Information Processing Systems*, Lake Tahoe, NV, USA, December 2012.
- [32] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 2017.
- [33] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-cam: visual explanations from deep networks via gradient-based localization,” in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 2017.
- [34] L. Prechelt, *Early Stopping-but when? Neural networks: Tricks of the Trade*, Springer, 1998.
- [35] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, July 2016.
- [36] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015.