STEMedicine

Alcoholic brain injury via 8-layer deep convolutional neural network

Ziquan ZHU[1,*], Mackenzie BROWN[2,*]

1 Department of Civil Engineering, University of Florida, Gainesville, United States
2 School of Engineering, Edith Cowan University, Joondalup WA 6027, Australia

*Correspondence: Zhu.ziquan@ufl.edu, mackbrown@ieee.org
https://doi.org/10.37175/stemedicine.v2i8.97

ABSTRACT

Alcohol can act quickly in the human body and alter mood and behavior. If people drink too much alcohol, it will accumulate in the liver and brain. To a certain extent, the symptoms of alcoholism will appear. So far, the main method of diagnosis of alcoholic brain injury is through MRI images by radiologists. However, this is a very subjective diagnosis. Radiologists could be influenced by external factors, resulting in diagnostic errors, such as physical discomfort, inattention, lack of rest, etc.. In the paper, we introduced a 8-layer deep convolutional neural network structure for alcoholic brain injury detection. Three fully connected layers, five pooling layers, and five convolution layers formed the proposed neural network structure. We proposed three improvements in this paper, (i) An automatic diagnosis method of alcoholic brain injury based on deep learning was proposed; (ii) We introduced Dropout to the proposed structure to improve robustness; (iii) Compared with other seven advanced approaches, the proposed 8-layer deep convolutional neural network structure is more efficient. The experimental results demonstrated that the specificity, sensitivity, accuracy, precision, FMI MCC and F1 were 96.20 ± 1.47, 96.14 ± 1.99, 96.17 ± 1.55, 95.98 ± 1.54, 96.06 ± 1.62 93.34 ± 3.11, 96.05 ± 1.62, respectively. Based on the comparison results, our method had the excellent performance. The proposed method can be used as one of the methods to detect alcoholic brain injury based on MRI images.

Keywords: Alcoholism · Deep convolution neural network · Dropout · MRI

Introduction

Alcohol, which is present in different proportions in various kinds of alcohol, can act quickly in the human body and alter mood and behavior. It can spread and distribute directly into the bloodstream throughout the body. The liver is where alcohol is mainly metabolized in the body. After entering the human body, a small amount of alcohol can be immediately discharged through the lungs or sweat glands. In the liver, the vast majority of alcohol reacts with alcohol dehydrogenase to produce acetaldehyde. People cannot metabolize alcohol indefinitely. If people drink too much alcohol, it

would accumulate in the liver and brain. When alcohol accumulates in the human body to a certain extent, it will lead to alcoholism. In medicine, alcoholism is medically divided into chronic and acute alcoholism. Acute alcoholism is the central nervous system dysfunction caused by heavy consumption of alcohol or alcoholic beverages over a period of time, which is usually manifested as abnormal behavior and consciousness. Severe damage to organ function can result in respiratory and circulatory failure, and then endanger life. Chronic alcoholism is a serious poisoning of the central nervous system caused by excessive drinking over a long period of time. The human body will not absorb and digest ethanol in alcohol. The ethanol in alcohol flows into the brain through the circulation of blood, further damages the nerve cells in the human brain, causes harm to the central nervous system, leads to the increase of inhibitory neurons, inhibit active neurons, and makes people slow to

respond. Once the excessive accumulation of ethanol in alcohol will cause irreparable damage to nerve cells. Patients with chronic alcoholism encephalopathy may have Wernicke's encephalopathy, Kosakoff syndrome, chronic alcoholic dementia, alcoholic delirium, alcoholic epilepsy, alcoholic psychotropic behavior disorder, and so on. Excessive drinking can also cause an alcoholic coma. Alcoholism is very common in some countries (e.g., Europe, America). Alcoholism is the third leading public health issue in the United States, after cardiovascular disease and cancer.

With the continuous research and development of deep learning, deep learning technology has been widely used in many fields, such as medicine. Lu proposed a radial-basis-function neural network (RBFNN) to detect pathological brains (1). S Sudha used convolutional neural network (CNN) for automatic segmentation of RoI in medical imaging modalities such as carotid artery ultrasound images (2). Y.-C. Hsieh used VGG16 to find out breast MC clusters from the image. Moreover, they used Mask RCNN to find MCs from the clusters to remove the noise from the background (3). S. Kido developed an image-based computer-aided detection (CADe) algorithm by use of regions with CNN features (R-CNN) for detection of lung abnormalities (4). K. Karthik used benchmark deep learning models to improve the quality of diagnostic images, through Super-resolution, for enabling faster and easier detection of anomalies that may be missed otherwise (5). Y, Bar explored the ability of CNN learned from a non-medical dataset to identify different types of pathologies in chest x-rays (6). The best performance was achieved using CNN and GIST features. C.-Y. Sun proposed a novel deformable cascade faster region-based convolutional neural networks (Faster R-CNN) to realize vascular plaque recognition in OCT images (7). W. S. Kim presented a new automatic CAC voxel classification model with multi-scale CNN architecture which can reflect the advantages of large receptive CNN and small receptive CNN (8). W. Li et al. demonstrated a new region-based convolutional neural network framework for multi-task prediction using an epithelial network head and a grading network head (9). Compared with a single-task model, their multi-task model can provide complementary contextual information, which contributes to better performance. L. Li, J introduced a new convolutional neural network structure for medical image denoising-deep neural network based on wavelet domain (deep wavelet denoising net, DWDN) (10).

In the field of AI applied to alcoholism, we have collected results from existing publications. Hou used Hu moment invariant (HMI) to identify alcoholism (11). Yang combined HMI with machine learning (ML) to identify pathological brains (12). Their method is used as a comparison basis in this study. Han proposed a Three-Segment Encoded Jaya (TSEJ) to identify alcoholism (13). Chen, et al. employed fractional Fourier transform (FrFT) to identify pathological brains (14). Their method is used as a comparison method. Qian used cat swarm optimization (CSO) to identify alcoholism (15).

Macdonald employed logistic regression (LR) to identify alcoholism (16). Chen used wavelet energy entropy (WEE) to identify alcoholism (17). Xie presented an AlexNet transfer learning model to identifyalco holism (18).

So far, the main method of diagnosis of alcoholic brain injury is through MRI images by radiologists. However, this is a very subjective diagnosis. Radiologists could be influenced by external factors, resulting in diagnostic errors, such as physical discomfort, inattention, lack of rest, etc.. This paper proposed 8-layer customized deep convolution neural network to address this issue. There arethree improvements in this paper, (i) An automatic diagnosis method of alcoholic brain injury based on deep learning was proposed; (ii) We introduced Dropout to the proposed structure to improve stability and robustness; (iii) Compared with other seven advanced approaches, the proposed 8-layer deep convolutional neural network structure is more efficient.

The structure of paper is as follows: Section 2 introduces the materials, Section 3 talks about the basics of deep convolution neural network (DCNN), Section 4 describes the experimental and comparison results, conclusion in included in Section 5.

Dataset
Subjects

We strictly checked and excluded the data of the research subjects. Participants were first interviewed by computer diagnosis. We would exclude participants with serious mental illness. We would exclude participants who are not in a coma for more than 15 minutes caused by schizophrenia, depression, and alcoholism. If participants suffered from stroke, epilepsy, AIDS and Wernicke-Korsak of syndrome, we excluded them. Finally, the first language of the participants was mandarin, which was also one of requirements.

Participants can participate in the experiment through Nanjing Brain Hospital and Jiangsu Provincial People's Hospital or by Internet advertisements and flyer. We spent three years collecting data from 235 research subjects. All the data in this paper were agreed by the hospitals and participants. There were 121 non-alcoholic control participants in the study, including 59 males and 62 females. In addition, we collected 114 chronic alcohol participants (males-58, females-56) who had long-term abstinence.

Scan

The scanning instrument in this paper was Siemens VERIO TIM 3.0T Magnetic Resonance Scanner. During the scan, 235 participants kept calm and laid down with their eyes closed. MP-RAGE sequence was used to get a total of 216 transection slices of the participants' brain. The imaging parameters were TI = 900 ms, gray level depth = 8-bit, TR = 2000 ms, TE = 2.50 ms, FA = 9 °, FOV = 256 mm × 256 mm, matrix = 256 × 256. It is important to note here that in this paper our images replaced 16-bit
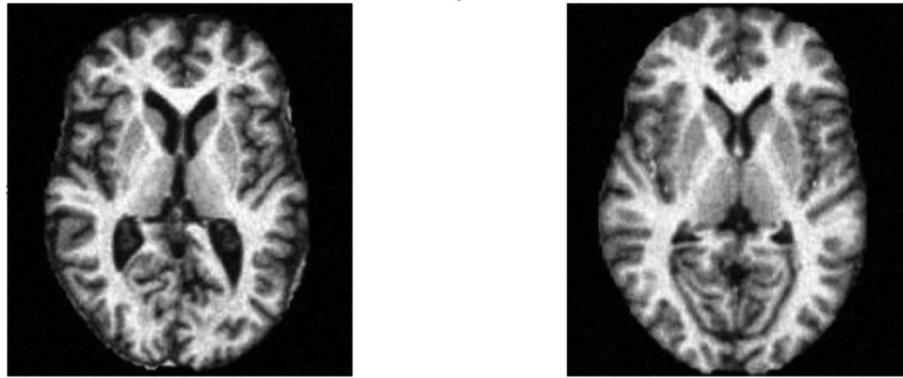
Figure 1. Slice example.

grayscale depth with 8-bit grayscale depth. Alcohol doesn't affect the depth of the brain images, so 8-bit grayscale depth was used in this paper. At the same time, the 8-bit grayscale depth of the images are enough to complete the experiment.

Slice Selection

FMRIB Software Library (FSL) V5.0 software was used to remove the skull and extract the brain. Standard MINI templates were standard for all volumetric images. Then each image is redefined as a 2 mm isotropic voxel. Finally, we selected the slice at G = 80 mm at MNI 152 coordinates, which was obtained linearly by 152-weighted MRI scanning plane, as shown in **Figure 1**. This location was chosen because it is the optimal location of the lateral ventricle and basal ganglia region. Finally we used a matrix of size 176 × 176 for subsequent classification.

Methodology

For the automatic diagnosis of alcoholic brain injury, we used a new method based on deep convolutional neural network (DCNN). Since 2012, deep convolutional neural network has been presented. More and more deep convolution neural network structures with better performance have been designed such as Alexnet (19), ResNet (20), Unet (21), VGGnet (22), and DenseNet (23). Due to the great progress of DCNN, DCNN has been applied in many fields, such as human detection (24), ear detection (25), etc. DCNN makes the network model no longer complex through three methods: local receptive field, weight sharing, and downsampling. Meanwhile, the depth convolution network structure has no effect on translation, rotation and scaling. In general, the common DCNN is composed of input layer, convolution layers, fully connected layers and output layer, as shown in **Figure 2**.

Figure 2 is a simple DCNN model, which contains three convolution layers, three pooling layers and three pooling layers. Finally, the results are exported through two fully connected layers.

Convolution layer

Convolution layer is the key parts of deep convolution neural network. Convolution layer uses convolution kernel to extract features (26, 27). Convolution kernel is a matrix (28). Starting from the upper left corner of the input matrix, the convolution kernel corresponds to the input matrix range, and then multiplies and adds to get a value (29). In this order, convolution is performed from left to right. This is the principle of convolution layer (30). The convolution layer formula is as follows:

$$W_{o+1} = \frac{W_o - K_w + 2B}{L} + 1 \qquad (1)$$

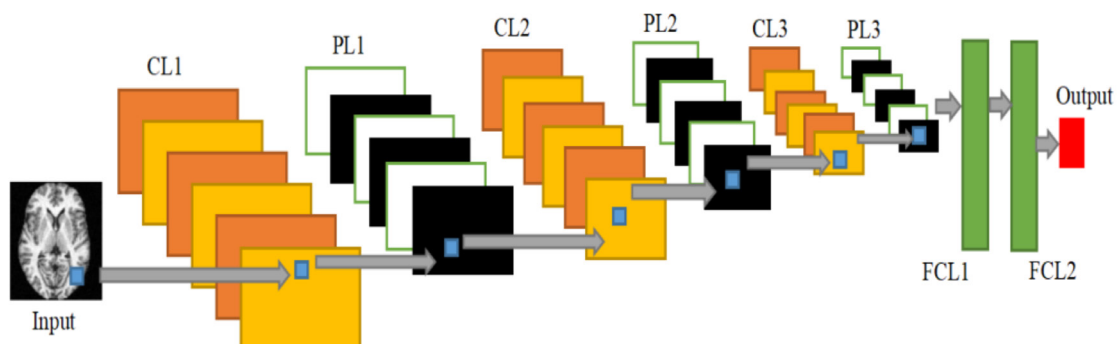$$H_{o+1} = \frac{H_o - K_h + 2B}{L} + 1 \qquad (2)$$
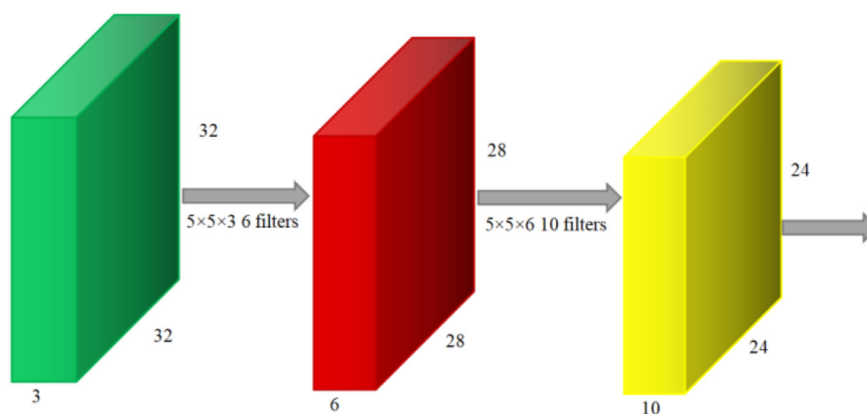

Figure 2. DCNN mo.

Figure 3. Convolution operation.

$$Do + 1 = N \qquad (3)$$

In the above formula, width is represented by W, height is represented by H, and depth is represented by D, the matrix of input is $W_o \times H_o \times D_o$, the matrix of output is $W_{o+1} \times H_{o+1} \times D_{o+1}$, K_w is the kernel width, K_h represents the kernel height, B is padding, L is stride, N is the kernel width, K_h represents the kernel height, B is padding, L is stride, N is the filters numbers.

Figure 3 is an example of image convolution. The input matrix is $32 \times 32 \times 3$, and 3 is its depth. The first convolution layer is composed of 6 filters with $5 \times 5 \times 3$. After the first convolution, the output matrix is $28 \times 28 \times 6$. The second convolution layer is composed of 10 filters with $5 \times 5 \times 6$. After the second convolution, the output matrix is $24 \times 24 \times 10$.

Figure 4 shows the flow chart of convolution layer.

Pooling layer

The feature map is obtained through convolution layer, and then these features are synthesized and classified. Theoretically, the features extracted from the convolution layer can be used as the input of the classifier. However, too many input values will cause a lot of calculation (31). At this time, the pooling layer is used to reduce the dimension of the feature map (32). There are two advantages for pooling layers: (i) it can make the feature graph smaller, simplify the calculation complexity of the network, reduce the parameters and calculations, and prevent over fitting (33); (ii) the receptive field can be increased by compressing features, extracting features and retaining main features, and keeping scale invariance.

Two methods are usually used for pooling layer, one is maximum pooling and another one is average pooling. Max pooling calculates the maximum value of the location and its adjacent matrix area and takes this maximum value as the value of the location (34). Generally, the average value of adjacent areas is calculated, and the calculated value is taken as the output value. This method is called average pooling. The pooling layer will not have any effect on depth of data matrix, but it can reduce the width and height to accomplish the goal of dimensionality reduction. Specific formula of pooling layer as follows:

$$MaxP = Max(E_C) \qquad (4)$$

$$AverP = \frac{\sum E_C}{|E_C|} \qquad (5)$$

As the formula above shows, E is the pooling region, C represents the activation set in the pooling region, E_C is elements numbers in the activation set.

Figure 5 shows the different results of different pooling methods. Taking the top left region as an example, the pooling stride of the two pooling methods is 2. The calculation process is as follows: AP = (1+3+5+8)/4 = 4.25; MP = max(1, 3, 5, 8) = 8.

Batch Normalization

For shallow models, data standardization preprocessing will be effective. However, With the continuous training of the model, the parameters of each layer are constantly changing. The closer to the output layer, the less stable the parameters would remain. When we train the deep neural network, although we have standardized the input data before training, the continuous updating of the model will lead to the instability of the output layer (35). Because of the instability of numerical, it is not easy to train a good deep neural network structure with great depth. Batch normalization (BN) continuously adjusts the neural network through the mean and standard deviation of small batches (36). Through this method, the value of each layer of the whole neural network can not fluctuate too much. The structure of batch normalization is shown as **Figure 6**.

Suppose u is the input, W is the weight and b is the deviation parameter. The output is x:

$$x = Wu + b \qquad (6)$$

Small batch B composed of m samples, B = {x(1), x(2). . . . x(n)}. Calculate the mean and variance of small batch B:

$$\mu_B \leftarrow \frac{1}{n} \sum_{i=1}^{n} x_i \qquad (7)$$

$$\sigma_B^2 \leftarrow \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_B)^2 \qquad (8)$$

The standardization for x_i:

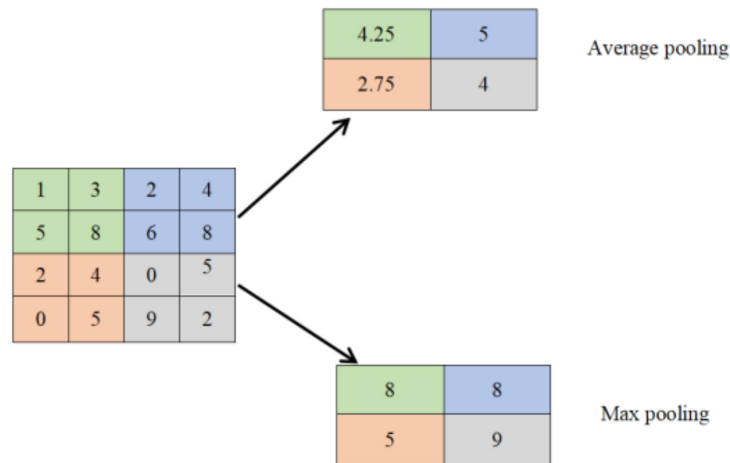$$\overline{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \Delta}} \qquad (9)$$

Figure 5. A example of average and max pooling.

This Δ > 0 is which Δ is a very small number to ensure that the denominator is greater than 0, which makes the formula and the results meaningful. Introduce learnable parameter: scale parameter φ And shift parameters ω.

$$y_i \leftarrow \varphi \overline{x}_i + \omega \qquad (10)$$

Dropout

In the deep learning structure, if the parameters of the structure are too many and the training samples are too few, over fitting is a common problem in trained models (37). The phenomenon of over fitting could appear in neural network training, which is shown in the following aspects (38): the structure has small loss function and high prediction accuracy in the training data. However, contrary to the results obtained from the training data, the model has large loss function and low prediction accuracy. If the model is over fitted, the resulting model can hardly be used (39). Dropout can be one of the solutions to reduce over fitting. In each training, by neglecting half of the feature detectors. It can significantly alleviate the problem of over fitting. The interaction between feature detectors (hidden layer nodes) can be reduced through dropout. **Figure 7** shows the difference between using dropout and not using dropout in deep learning.

Rectified Linear Unit

In the neuron, after the input is summed by weighting, the activation function is applied, as shown in **Figure 8**. The operation of the activation function is to activate some neurons in the neural network and transmit the activation information to the next layer of neural network (40). In order to increase neural network model nonlinearity, activation function is added. When there is no activation function, the output value of each layer in the convolutional neural network structure is a linear function (41). The output is the linear function of the input, which is the most primitive perceptron (42). In order to make the deep convolution neural network closer to the nonlinear function, a nonlinear factor is introduced into the activation function. In this way, deep neural network is suitable for a lot of nonlinear models (43). The activation function can also construct a sparse matrix to eliminate the redundancy in the data and retain the characteristics of the data as much as possible. The most commonly used activation functions are sigmoid and relu function, as shown in **Figure 9**. The formula of relu and sigmoid function:

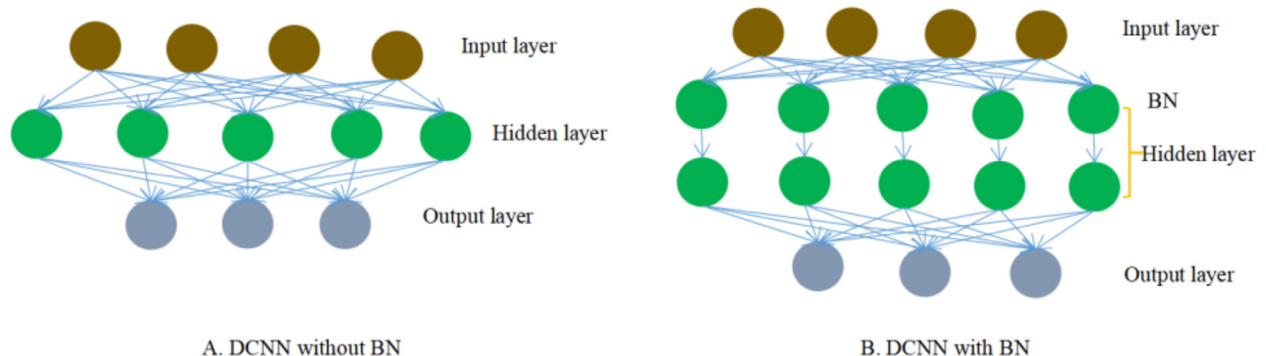$$relu = \begin{cases} 0, & \text{if } x \leq 0 \\ x, & \text{if } x > 0 \end{cases} \qquad (11)$$
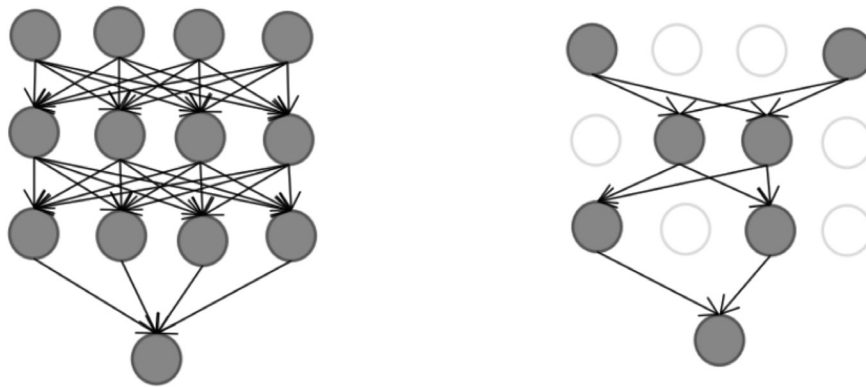

Figure 6. Structure of BN.

Figure 7. An example of network with dropout.

Comparing relu and sigmoid functions, (i) the sigmoid function has a problem of gradient vanishing. In contrast, the output of relu function is very stable, because it is a linear function when x > 0, and the problem of gradient disappearance is avoided; (ii) the calculation of sigmoid function is large; using relu function can save a large number of computation; (iii) the output of some neurons of relu function could be zero, which reduces the relationship between parameters and the density of the network, thus reducing the possibility of over fitting. Based on the above advantages, we used the relu function in this paper.

Structure of customized DCNN

In the paper, we introduced a 8-layer deep convolutional neural network structure for alcoholic brain injury detection. Three fully connected layers, five pooling layers, and five convolution layers formed the proposed neural network structure. Each convolution layer is connected to the pooling layer, and dropout is in front of each fully connected layer. The structure of 8-layer DCNN is shown in **Figure 10**.

The parameter settings of 8-layer DCNN structure are shown in **Table 1**. The input size is 176×176. The filter size of each convolution layer is 3×3. The number of filters in the first convolution layer is 32, that in the second convolution layer is 64, that in the third convolution layer is 128, that in the forth convolution layer is 256, and that in the fifth convolution layer is 512. After five times convolution and pooling, three dropout and three fully connected layers are added, and the final output size is 2.

Measures

For the evaluation index of binary classification, the results are divided into positive and negative categories. In the actual classification, there are four categories. Suppose the confusion matrix is:

$$\text{confusion matrix} = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix} \quad (13)$$

TP represents True Positive; FN is False Positive; FN represents False Negative; TN is True Negative. In the paper, seven indicators are used to evaluate the model.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (14)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (15)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (16)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (17)$$

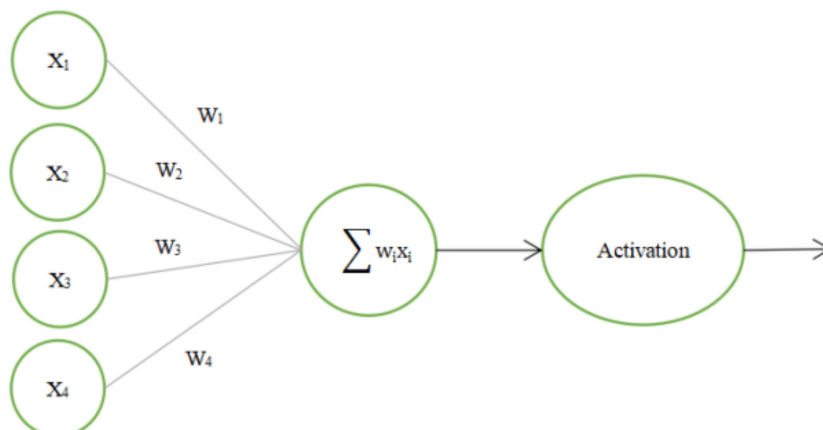$$F1 = \frac{2TP}{2TP + FP + FN} \quad (18)$$

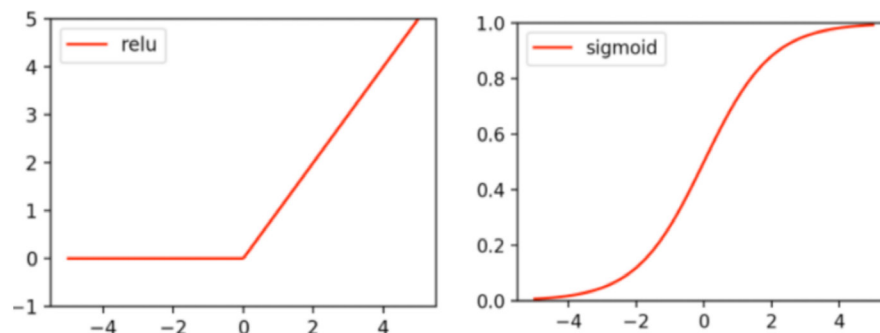**Figure 8. Schematic diagram of activation function.**

Figure 9. relu and sigmoid function.

$$MCC = \frac{TP \times TN - TP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \qquad (19)$$

$$FMI = \frac{TP}{\sqrt{(TP + FP) \times (TP + FN)}} \qquad (20)$$

In order to avoid contingency, 10 runs were carried out, and the mean and variance of each indicators were calculated.

Experiment Result
Statistical Analysis

In this paper, we ran 10 times and got ten groups of data. Each time the division was updated randomly. The experimental results are shown in **Table 2**. The average of specificity, sensitivity, accuracy, precision, FMI MCC and F1 were 96.20 ± 1.47, 96.14 ± 1.99, 96.17 ± 1.55, 95.98 ± 1.54, 96.06 ± 1.62 93.34 ± 3.11, 96.05 ± 1.62, respectively. From **Figure 11**, we can get this conclusion that the eighth group is the best and the sixth group is the eighth group is the best and the sixth group is the worst.

Comparison with State-of-the-art Approaches

We compare the proposed "8-layer customized DCNN"

with seven advanced approaches: HMI (11), HMI+ML (12), TSEJ (13), FrFT (14), CSO (15), LR (16), WEE (17). We have described these advanced approaches in detail in the first chapter introduction. Table 3 shows the comparison test results. The accuracy of the proposed DCNN structure is 96.17 ± 1.55. Our accuracy is at least 2% higher than that of other seven approaches. HMI (11) yields 89.15, HMI + ML (12) obtains 85.15, TSEJ (13) gets 93.66, FrFT [14] achieves 86.77, CSO (15) obtains 92.13, LR (16) gets 83.70 ± 1.35, WEE (17) achieves 92.55 ± 0.67. It is easy to conclude that the DCNN structure which is introduced in the paper has superior performance than these seven approaches. For clear view, a comparison plot is shown in **Figure 12**.

Conclusion

In the paper, we introduced a 8-layer deep convolutional neural network structure for alcoholic brain injury detection. Three fully connected layers, five pooling layers, and five convolution layers formed the proposed neural network structure. In this framework, we added dropout to reduce over fitting problem. We compared our

Table 1. Parameter settings of deep convolutional neural network.

Layers	Size	Parameters
Input	176 × 176	
Conv Layer -first	88 × 88 × 32	32 3 × 3, pooling size = 2
Conv Layer -second	44 × 44 × 64	64 3 × 3, pooling size = 2
Conv Layer -third	22 × 22 × 128	128 3 × 3, pooling size = 2
Conv Layer -forth	11 × 11 × 256	256 3 × 3, pooling size = 2
Conv Layer -fifth	5 × 5 × 512	512 3 × 3, pooling size = 2
Dropout-1		0.5
FCL-1	1000	12800 × 1000
Dropout-2		0.5
FCL-2	100	1000 × 100
Dropout-3		0.5
FCL-3	2	100 × 2

Figure 10. Structure of DCNN.

Table 2. Experimental result.

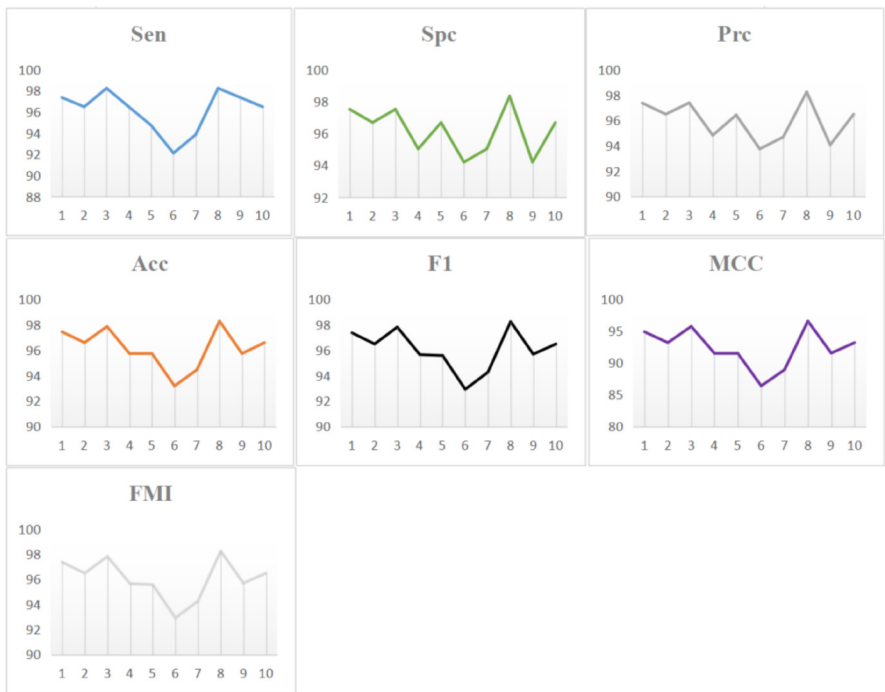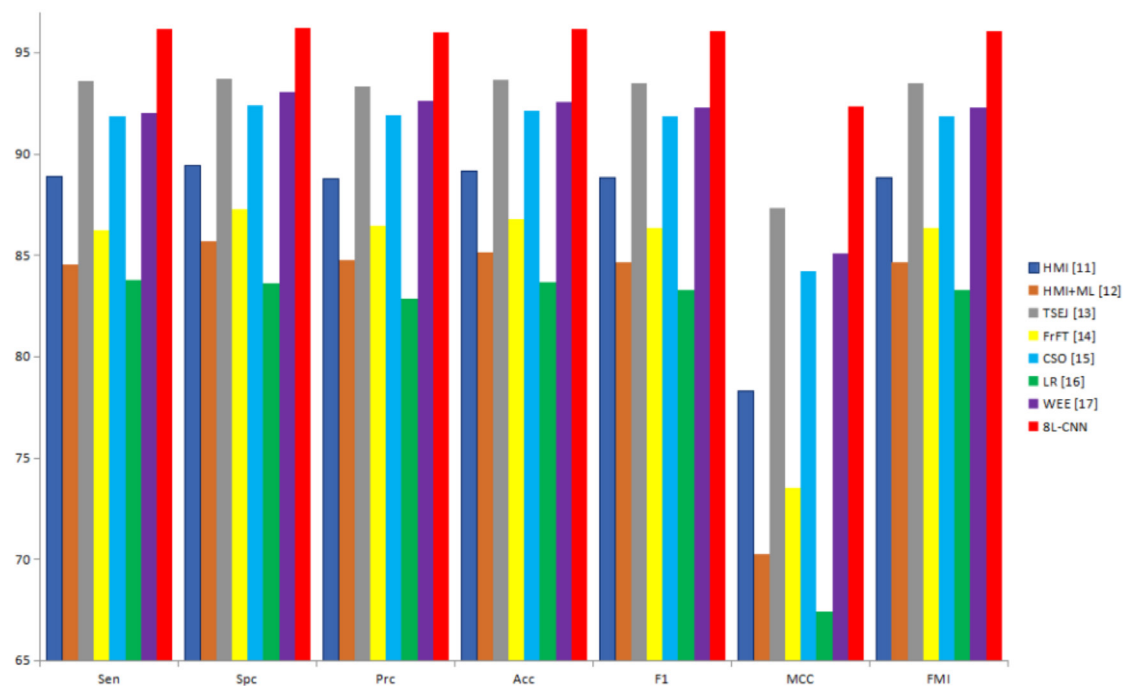| Run	sensitivity	specificity	precision	accuracy	F1	MCC	FMI
1	97.37	97.52	97.37	97.45	97.37	94.89	97.37
2	96.49	96.69	96.49	96.60	96.49	93.19	96.49
3	98.25	97.52	97.39	97.87	97.82	95.75	97.82
4	96.49	95.04	94.83	95.74	95.65	91.50	95.66
5	94.74	96.69	96.43	95.74	95.58	91.49	95.58
6	92.11	94.21	93.75	93.19	92.92	86.38	92.92
7	93.86	95.04	94.69	94.47	94.27	88.93	94.27
8	98.25	98.35	98.25	98.30	98.25	96.59	98.25
9	97.37	94.21	94.07	95.74	95.69	91.54	95.70
10	96.49	96.69	96.49	96.60	96.49	93.19	96.49
MSD	96.14 ± 1.99	96.20 ± 1.47	95.98 ± 1.54	95.98 ± 1.54	96.05 ± 1.62	92.34 ± 3.11	96.06 ± 1.62

Table 2. Experimental result.

Figure 12. comparison of seven state-of-the-art approache.

method with other seven advanced approaches to test the performance of the DCNN structure. According to the comparison test results, we validated the DCNN structure had the excellent performance. The proposed model can be used as one of the methods to detect alcoholic brain injury based on MRI images.

The limitations of the paper: (i) there are very few training data set. With the number of cases increases and the amount of training increases, the performance of the system will achieve higher accuracy; (ii) we did not compare DCNN structures with different number of hidden layers.

In the next paper, more data will be collected in our next experiment to test performance. Meanwhile, getting the best combination of hidden layers is also one of the main purposes of our next paper. What's more, some new technologies may be applied to our next paper for comparison, such as transfer learning pre-trained models.

Conflict of interest
The authors declare that they have no conflicts of interest to disclose.

References
1. Lu Z, Lu S, Liu G, Zhang Y, Yang J, Phillips P. A pathological brain detection system based on radial basis function neural network. J Med Imaging Health Inform. 2016;6:1218-1222.
2. Sudha S, Jayanthi KB, Rajasekaran C, Sunder T, editors. Segmentation of RoI in medical images using CNN- a comparative study. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON); 2019 17-20 Oct. 2019.
3. Hsieh YC, Chin CL, Wei CS, Chen IM, Yeh PY, Tseng RJ, editors. Combining VGG16, Mask R-CNN and Inception V3 to identify the benign and malignant of breast microcalcification clusters. 2020 International Conference on Fuzzy Theory and Its Applications (iFUZZY); 2020 4-7 Nov. 2020.
4. Kido S, Hirano Y, Hashimoto N, editors. Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN). 2018 International Workshop on Advanced Image Technology (IWAIT); 2018 7-9 Jan. 2018.

Table 3. The comparison results.

Method	Sen	Spc	Prc	Acc	F1	MCC	FMI
HMI [11]	88.86	89.42	88.78	89.15	88.82	78.28	88.82
HMI+ML [12]	84.56	85.70	84.78	85.15	84.67	70.27	84.67
TSEJ [13]	93.60	93.72	93.35	93.66	93.47	87.31	93.47
FrFT [14]	86.23	87.27	86.46	86.77	86.34	73.51	86.34
CSO [15]	91.84	92.40	91.92	92.13	91.88	84.24	91.88
LR [16]	83.77 ± 2.42	83.64 ± 2.33	82.88 ± 1.90	83.70 ± 1.35	83.29 ± 1.42	67.44 ± 2.66	83.31 ± 1.41
WEE [17]	92.02 ± 1.13	93.06 ± 1.42	92.61 ± 1.36	92.55 ± 0.67	92.30 ± 0.67	85.11 ± 1.35	92.31 ± 0.68
8L-CNN (Ours)	96.14 ± 1.99	96.20 ± 1.47	95.98 ± 1.54	96.17 ± 1.55	96.05 ± 1.62	92.34 ± 3.11	96.06 ± 1.62

5. Karthik K, Kamath SS, Kamath SU, editors. Automatic quality enhancement of medical diagnostic scans with deep neural image super-resolution models. 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS); 2020 26-28 Nov. 2020.

6. Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, Greenspan H, editors. Chest pathology detection using deep learning with non-medical training. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI); 2015 16-19 April 2015.

7. Sun CY, Hong XJ, Shi S, Shen ZY, Zhang HD, Zhou LX. Cascade Faster R-CNN Detection for Vulnerable Plaques in OCT Images. IEEE Access. 2021;9:24697-704.

8. Kim WS, Jung HY, Choi JH, editors. Automatic CAC voxel classification with multi-scale CNN architecture. 2019 International Conference on Information and Communication Technology Convergence (ICTC); 2019 16-18 Oct. 2019.

9. Li W, Li J, Sarma KV, Ho KC, Shen S, Knudsen BS, et al. Path R-CNN for prostate cancer diagnosis and gleason grading of histological images. IEEE Trans Med Imaging. 2019;38(4):945-54.

10. Li L, Wu J, Jin X. CNN denoising for medical image based on wavelet domain. in 2018 9th International Conference on Information Technology in Medicine and Education (ITME), Hangzhou, China, 2018;pp 105-109.

11. Zhang Y-D, Zhang Y, Lv Y-D, Hou X-X, Liu F-Y, Jia W-J, et al. Alcoholism detection by medical robots based on Hu moment invariants and predator–prey adaptive-inertia chaotic particle swarm optimization. Comput Electr Eng. 2017;63:126-138.

12. Zhang Y, Yang J, Wang S, Dong Z, Phillips P. Pathological brain detection in MRI scanning via Hu moment invariants and machine learning. J Exp Theor Artif Intell. 2016;29:299-312.

13. Wang S-H, Muhammad K, Lv Y, Sui Y, Han L, Zhang Y-D. Identification of alcoholism based on wavelet Renyi entropy and three-segment encoded Jaya algorithm. Complexity. 2018;2018:3198184.

14. Zhang Y-D, Chen S, Wang S-H, Yang J-F, Phillips P. Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine. Int J Imaging Syst Technol. 2015;25;317-327.

15. Zhang Y-D, Sui Y, Sun J, Zhao G, Qian P. Cat Swarm Optimization applied to alcohol use disorder identification. Multimed Tools Appl. 2018;77:22875-22896.

16. Yuanyuan T, Felix M, editors. Alcoholism Detection via Wavelet Energy and Logistic Regression. Proceedings of the 2018 International Conference on Information Technology and Management Engineering (ICITME 2018); 2018 2018/08: Atlantis Press.

17. Chen X and Yan Y. Alcoholism detection by wavelet energy entropy and linear regression classifier. Cmes-comp Model Eng. 2021;127:325--343.

18. Wang S-H, Xie S, Chen X, Guttery DS, Tang C, Sun J, et al. Alcoholism identification based on an AlexNet transfer learning model. Front Psychiatry. 2019;10:205.

19. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60:84-90.

20. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016 pp. 770-778.

21. Ronneberger O, Fischer P, Brox T, editors. U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; 2015 2015//; Cham: Springer International Publishing.

22. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015 pp. 3431-3440.

23. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017 pp. 2261-2269.

24. Kim JH, Hong HG, Park KR. Convolutional neural network-based human detection in nighttime images using visible light camera sensors. Sensors (Basel). 2017;17(5):1065.

25. Cintas C, Quinto-Sánchez M, Acuña V, Paschetta C, de Azevedo S, Cesar Silva de Cerqueira C, et al. Automatic ear detection and feature extraction using Geometric Morphometrics and convolutional neural networks. IET Biometrics. 2017;6:211-223.

26. Hwang C, Kim D, Lee T. Semi-supervised based unknown attack detection in EDR environment. KSII Trans Internet Inf Syst. 2020;14:4909-4926.

27. Jung H and Lee BG. The impact of transforming unstructured data into structured data on a churn prediction model for loan customers. KSII Trans Internet Inf Syst. 2020;14:4706-4724.

28. Wang S-H, Wu K, Chu T, Fernandes SL, Zhou Q, Zhang Y-D, et al. SOSPCNN: structurally optimized stochastic pooling convolutional neural network for tetralogy of fallot recognition. Wirel Commun Mob Comput. 2021;2021:5792975.

29. Altwaijry N. Keystroke dynamics analysis for user authentication using a deep learning approach. Int J Comput Sci Inf Secur. 2020;20:209-216.

30. Zhang Y-D, Satapathy SC, Wu D, Guttery DS, Górriz JM, Wang S-H. Improving ductal carcinoma in situ classification by convolutional neural network with exponential linear unit and rank-based weighted pooling. Complex Intell Syst. 2021;7:1295-1310.

31. Afrasiabi S, Afrasiabi M, Mohammadi M, Parang B. Fault localisation and diagnosis in transmission networks based on robust deep Gabor convolutional neural network and PMU measurements. IET Gener Transm Distrib. 2020;14:6484-6492.

32. Wang S-H, Zhou Q, Yang M, Zhang Y-D. ADVIAN: alzheimer's disease VGG-inspired attention network based on convolutional block attention module and multiple way data augmentation. Front Aging Neurosci. 2021;13:687456.

33. Sahani M, Swain B, Dash P. FPGA-based favourite skin colour restoration using improved histogram equalization with variable enhancement degree and ensemble extreme learning machine. IET Image Process. 2021;15:1247-1259.

34. Zhang Y, Zhang X, Zhu W. ANC: attention network for COVID-19 explainable diagnosis based on convolutional block attention module. Comput Model Eng Sci. 2021;127:1037-1058.

35. Suzuki H, Yoshitaka T, Yoshio TA-O, Tada T. Artificial intelligence for cancer detection of the upper gastrointestinal tract. Dig Endosc. 2021;33:254-262.

36. Wang S-H, Zhang Y, Cheng X, Zhang X, Zhang Y-D. PSSPNN: PatchShuffle stochastic pooling neural network for an explainable diagnosis of COVID-19 with multiple-way data augmentation. Comput Math Methods Med. 2021;2021:6633755.

37. Trasolini R, Byrne MF. Artificial intelligence and deep learning for small bowel capsule endoscopy. Dig Endosc. 2021;33:290-297.

38. Hirasawa T, Ikenoyama Y, Ishioka M, Namikawa K, Horiuchi Y, Nakashima H, et al. Current status and future perspective of artificial intelligence applications in endoscopic diagnosis and management of gastric cancer. Dig Endosc. 2021;33:263-272.

39. Zhang Y-D, Satapathy SC, Guttery DS, Górriz JM, Wang S-H. Improved breast cancer classification through combining graph convolutional network and convolutional neural network. Inf Process Manag. 2021;58:102439.

40. Corazza PA-O, Maddison J, Bonetti P, Guo L, Luong V, Garfinkel A, et al. Predicting wet age-related macular degeneration (AMD) using DARC (detecting apoptosing

retinal cells) AI (artificial intelligence) technology. Expert Rev Mol Diagn. 2021;21:109-118.

41. Zhang YD, Satapathy SC, Liu S, Li GR. A five-layer deep convolutional neural network with stochastic pooling for chest CT-based COVID-19 diagnosis. Mach Vis Appl. 2021;32:14.

42. Joshi S, Kumar R, Dwivedi A. Hybrid DSSCS and convolutional neural network for peripheral blood cell recognition system. IET Image Process. 2020;14:4450-4460.

43. Wang S-H, Govindaraj VV, Górriz JM, Zhang X, Zhang Y-D. Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. Inf Fusion. 2021;67:208-229.