Clinics of Oncology

Research Article ISSN: 2640-1037 | Volume 6

Image Classification of Brain Tumor Based on Enhanced VGG 19 Convolutional Neural Network

Tian C[#1], Xi Y[#2], Ma Y[2], Chen C[4], Wu C[2], Ru K[3], Li W[*1] and Zhao M[3*]

[1]School of Control Science and Engineering, Shandong University, Jinan, Shandong, 250061, China

[2]Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, Shandong, China

[3]Department of Pathology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, China

[4]Shandong Institute of Advanced Technology, Chinese Academy of Sciences, Jinan, Shandong, China

[*]Corresponding author:

Wei Li,
School of Control Science and Engineering, Shandong University, Jinan, Shandong, 250061, China
and Miaoqing Zhao, Shandong Institute of Advanced Technology, Chinese Academy of Sciences, Jinan, Shandong, China

Received: 10 Apr 2023
Accepted: 24 May 2023
Published: 02 June 2023
J Short Name: COO

Keywords:

Deep learning; Convolutional neural network; Primary central nervous system diffuse large B-cell lymphoma; High-grade glioma; VGG19

Citation:

Li W, Image Classification of Brain Tumor Based on Enhanced VGG 19 Convolutional Neural Network. Clin Onco. 2023; 6(23): 1-10

Abbreviations:

HGG: High-grade glioma; CNS-pDLBCL:Primary central nervous system diffuse large B-cell lymphoma; AI:Artificial intelligence; PCNSLs:Primary central nervous system lymphomas; CNS:Central nervous system; DLBCL:Diffuse large B cell lymphoma; CNN:Convolutional neural network; ROI:Region of interest; TP:True positive; FN:False negative; TN:True negative; FP:False positive; FCL:Fully connected layer

1. Abstract

The clinical symptoms and imaging studies of primary central nervous system diffuse large B-cell lymphoma (CNS-pDLBCL) and high-grade glioma (HGG) have no clear specificity, and pathology is the gold standard for diagnosis. However, there are issues such as complicated pathological slide preparation, significant subjectivity in diagnosis, and disease histological forms diversification. When the diagnostic result is uncertain, additional judgments utilizing immunohistochemistry are required, complicating the diagnosis work. To overcome the aforementioned challenges, we offer a novel way to assisting pathologists in their diagnostic job. An enhanced VGG19 convolutional neural network is used in this study. To improve classification stability and accuracy, we keep the original convolutional and pooling layers and replace the Softmax layer in the model with a support vector machine classifier, as well as incorporating the transfer learning method into the model. The ten-fold cross-validation method was employed during the training and testing procedure to increase the amount of data and improve the model's generalization capacity. Furthermore, when compared to existing models and deep learning methods, the utilized model outperforms them with a 95% accuracy rate. As a result, this model can be employed for pathologically objective, accurate, and quick diagnosis of CNS-pDLBCL and HGG.

2. Introduction

Diffuse large B-cell lymphoma (DLBCL) accounts for more than 95% of primary central nervous system lymphoma (PCNSL). Early detection is difficult due to the unique site of tumor growth and the variety of clinical symptoms, and patient survival is low [1]. High-grade glioma (HGG) is a poorly differentiated glioma that accounts for a relatively high proportion of brain gliomas. This is a malignant tumor with a terrible prognosis [2].

The prevalence of these diseases has been rising in recent years. Both feature high cell density and invasive growth of lesions, and the location of the disease is similar, making mistake and

missed diagnosis easy. Clinical diagnosis is mostly determined by symptoms and imaging features [3]. Imaging is a straightforward, convenient, and non-invasive method of examining brain illnesses. CT and MRI are currently the most routinely utilized imaging techniques in the diagnosis of the two malignancies [4].Subsequent research has revealed that it is easily confused when employing conventional imaging techniques [5-7]. Liu Rui et al [8] employed CT to detect lymphoma with an accuracy rate of 70%, and the accuracy rate for glioma identification was 75%; MRI detection accuracy rate for lymphoma was 82.5%, and the accuracy rate for glioma detection was 85%. Geng Lei et al [9] employed image detection technology to discriminate lymphoma and glioma with an accuracy percentage of 89.7%. In general, the accuracy of MRI and CT detection of the two tumors is not perfect, so it is critical to enhance diagnostic accuracy.

When imaging differences are difficult to discern, the generally employed clinical pathological diagnosis discloses their nature and serves as a foundation for further treatment. There is no universally accepted method for identifying PCNSL and HGG. Although they can be identified by pathology, they have different histological forms, and there are issues such as complicated pathological section preparation, excessively long times, and the requirement for immunohistochemistry auxiliary diagnostic [10]. Pathological diagnosis is strongly reliant on pathologists' professional expertise and diagnostic experience, and pathologists are in short supply in resource-constrained places [14]. Even the most skilled pathologists make mistakes. They frequently only give a most likely disease type without quantifying the answer, which may cause the patient's therapy to be delayed. As a result, effective diagnostic methods are required to improve the accuracy of preoperative diagnosis, and it is critical to find a reliable and objective diagnostic method to assist pathologists in diagnosis, solve the problem of uneven resource allocation, and provide accurate clinical treatment guidance.

Deep learning (DL) [11] has seen a lot of success in computer vision. Deep learning is a subfield of machine learning that consists of multilayer neural networks with an input layer, an output layer, and numerous hidden layers [12]. Convolutional neural network (CNN) is a deep feed-forward neural network that can analyze visual pictures and capture more features. Artificial intelligence techniques, particularly DL models, are being actively used in the development of computer-aided disease diagnosis tools [13]. Many CNN frameworks for image classification have been presented in recent years, ushering in a new era of manual feature extraction.

This study's primary contributions are as follows: Firstly, in light of the shortcomings in pathological diagnosis, we created and deployed a deep learning model suitable for the classification and identification of PCNSL and HGG. Secondly, when compared to the traditional network model, the deep network model we utilized showed higher accuracy and sensitivity, showing that our model can better recognize PCNSL and HGG. Finally, considering the scarcity of studies on deep learning in the classification and diagnosis of PCNSL and HGG disease, we might propose a new study direction.

3. Materials and Methods

3.1. Materials

3.1.1. CNN

AlexNet [15]won in the ImageNet Classification Competition in 2012 due to its superior classification effect compared with other models in the competition. Deep learning has attracted wide attentions since then. Compared with traditional machine learning, deep learning is superior because it does not require manual feature extraction. Instead, it learns the features of samples by using neural network and improves the convenience and accuracy of feature extraction. As one of the most common deep learning models, CNN shows excellent performance in image processing. In this study, CNN was applied to construct the image classification model. CNN is usually composed of an input layer, a convolutional layer, a pooling layer, and a fully connected layer (FCL) (Figure 1). The input is 2D or 3D images, the convolutional layer is used to extract the features of images, and the pooling layer reduces the dimension of extracted features and the quantities of data and parameters. After a series of convolutional and pooling operations, CNN can simultaneously learn the low-layer and high-layer characteristics of data. The FCL acquires the features easily distinguishable by the network and convenient for subsequent classification. Compared with traditional neural network, CNN has two significant advantages: local connection and weight sharing. Local connection is relative to full connection. Full connection means that each node in the network is connected, and local connection means that only some nodes are connected. In practical processing, the pixel points of images usually have a strong correlation with nearby pixel points but a weak correlation with the far ones. Local connection can form local features that are easy to be distinguished. Weight sharing refers to convolutional operation of the whole image by using the same convolutional kernel. It can decrease parameter size during operation and accelerate operation speed.

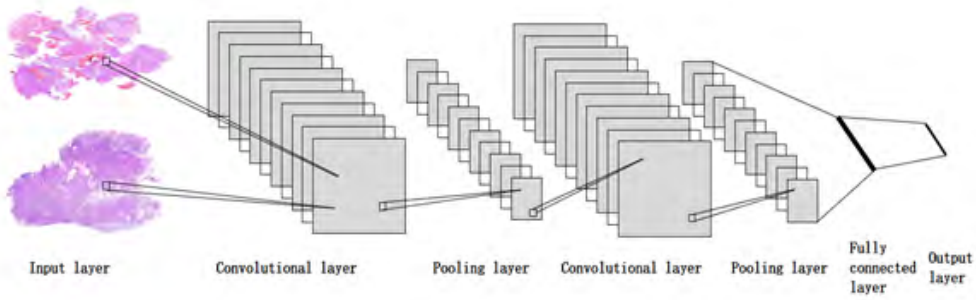

Figure 1: Typical structure of CNN

3.1.2. VGG19

Compared with LeNet[16], VGG[17] network has one improvement of replacing the large-sized convolutional kernel with several small-sized ones. For example, VGG uses two 3X3 convolutional kernels to replace the 5X5 convolutional kernel and uses three 3X3 convolutional kernels to replace the 7X7 convolutional kernel. This feature has one advantage: given the same reception field, stacking several small-sized convolutional layers can improve the depth of the network and increase the ability of feature extraction (increase the nonlinear layers). Moreover, this step involves few parameters. For example, a size 5 reception field is equivalent to stacking two 3X3 convolutional kernels with a step length of 1 (that is, a 5X5 convolutional kernel is equal to two 3X3 convolutional kernels). The parameter size of a 5X5 convolutional kernel is 5*5*C2, and that of two 3X3 convolutional kernels is 2*3*3*C2. Hence, 18C2 < 25C2, indicating that the 3X3 convolutional kernel is beneficial to preserve the image properties. Offset was ignored in the calculation of parameter size. The numbers of input and output channels were hypothesized to be C. The VGG19 network structure is shown in (Figure 2).

3.1.3. SVM

SVM [18] is a type of generalized linear classifiers used for binary data categorization in supervised learning. Corinna Cortes and Vapnik proposed it in 1993, and it was published in 1995. SVM is a sophisticated machine learning method that works well in a variety of applications. It is well suited to high-dimensional datasets due to its ability to handle high-dimensional features efficiently. Furthermore, SVM is extremely effective for small datasets because it does not require a big amount of training data. However, because SVM has a high computational complexity, it might become slow when dealing with huge datasets. Prior to the advent of deep learning (2012), SVM was regarded as the most successful and best-performing machine learning algorithm in the previous ten years. SVM can be applied to both classification and regression problems. It has the advantages of strong generalization performance, small sample size, and so on, and it is frequently employed in a variety of practical issues.

3.1.4. Transfer Learning

Transfer learning [19] is a novel machine learning learning paradigm that can tackle the problem of limited data sets in the field of medical image analysis, efficiently expedite model training, and prevent over-fitting issues. Transfer learning is currently being used in the realm of medical picture analysis. Migration learning is used to overcome the limitations of highly specialized but small-scale data sets. First, the convolutional layer of the pre-trained model is migrated to the target domain, and then the data of the target domain is used to retrain to adjust the parameters [20]. A model trained on a large-scale picture dataset, for example, can be applied to an image classification problem and fine-tuned on a new image classification challenge.

3.1.5. Image Processing

Deep learning algorithmic models outperform more standard machine learning approaches, but they require larger datasets for training to avoid overfitting [21]. To solve this issue, data augmentation has become a prominent method for increasing the amount of training datasets in medical image processing [22]. Shear transformation, flip transformation, scaling transformation, translation transformation, scale transformation, contrast transformation, noise perturbation, color transformation, and other regularly used picture data augmentation methods are listed below. Large datasets are required for deep neural network applications to succeed. Image augmentation expands the training set by producing similar but distinct training samples following a sequence of random adjustments to the training images. Furthermore, changing the training samples at random can lessen the model's reliance on specific traits, hence enhancing the model's generalization ability. For example, we crop photographs in various ways so that objects of interest appear in varied positions, minimizing the model's reliance on where objects appear. We can also change the brightness, hue, and other parameters to lessen the model's sensitivity to color. We discovered in past practice that as more image space improvement techniques are utilized, the model's accuracy decreases due to the problem of data specificity. As a result, we solely used flipping and translating space enhanced technology in this investigation, as well as image contrast and saturation adjustments.

3.1.6. Classification Model

We have a small sample size, thus the findings are not perfect whether using shallow models (SVM, LeNet, AlexNet) or deeper network models (VGG19, VGG16, Resnet [23]). Tang Y [24]

mostly replaced the Softmax classifier with the SVM classifier, and the cross-entropy loss was replaced with the hinge loss. To categorize pictures of bacterial and viral pneumonia, Xiong Feng [25] coupled the VGG19 convolutional neural network with machine learning methods to create two better VGG19 models based on SVM (linear) and XGBoost. Based on this, plus the fact that we previously solely utilized VGG19 for training, it is easy to induce overfitting and low accuracy. We also investigated the combination of machine learning algorithms and deep learning models, which we successfully applied to PCNSL and HGG picture categorization (as shown in Figure 3). In this way, the features learned by our model are not directly input into three fully connected layers like the original VGG19 and then input into the Softmax layer to finally get the classification result, but the method we use is that all the extracted features are directly input through the fully connected layer into the linear interface of the SVM, and then give the final classification result. Simultaneously, with the help of transfer learning, our training speed has increased, and the phenomenon of overfitting has been greatly decreased. Softmax primarily decentralizes and then classifies based on the key degree of feature vectors acquired through convolution, whereas SVM splits all features based on the distribution of feature vectors in space. In contrast, SVM will catch some subtle features and may make full use of each feature, resulting in a greater identification rate. At the same time, the calculation time complexity of SVM is lower than that of Softmax, which can save a significant amount of time.

<table>
<tr><td colspan="6">ConvNet Configuration</td></tr>
<tr><td>A</td><td>A-LRN</td><td>B</td><td>C</td><td>D</td><td>E</td></tr>
<tr><td>11 Weight Layers</td><td>11 Weight Layers</td><td>13 Weight Layers</td><td>16 Weight Layers</td><td>16 Weight Layers</td><td>19 Weight Layers</td></tr>
<tr><td colspan="6">Input(224*224 RGB Image)</td></tr>
<tr><td rowspan="2">conv3-64</td><td>conv3-64</td><td>conv3-64</td><td>conv3-64</td><td>conv3-64</td><td>conv3-64</td></tr>
<tr><td>LRN</td><td>conv3-64</td><td>conv3-64</td><td>conv3-64</td><td>conv3-64</td></tr>
<tr><td colspan="6">Maxpool</td></tr>
<tr><td rowspan="2">conv3-128</td><td rowspan="2">conv3-128</td><td>conv3-128</td><td>conv3-128</td><td>conv3-128</td><td>conv3-128</td></tr>
<tr><td>conv3-128</td><td>conv3-128</td><td>conv3-128</td><td>conv3-128</td></tr>
<tr><td colspan="6">Maxpool</td></tr>
<tr><td>conv3-256</td><td>conv3-256</td><td>conv3-256</td><td>conv3-256</td><td>conv3-256</td><td>conv3-256</td></tr>
<tr><td rowspan="3">conv3-256</td><td rowspan="3">conv3-256</td><td rowspan="3">conv3-256</td><td>conv3-256</td><td>conv3-256</td><td>conv3-256</td></tr>
<tr><td rowspan="2">conv1-256</td><td rowspan="2">conv3-256</td><td>conv3-256</td></tr>
<tr><td>conv3-256</td></tr>
<tr><td colspan="6">Maxpool</td></tr>
<tr><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td></tr>
<tr><td rowspan="3">conv3-512</td><td rowspan="3">conv3-512</td><td rowspan="3">conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td></tr>
<tr><td rowspan="2">conv3-512</td><td rowspan="2">conv3-512</td><td>conv3-512</td></tr>
<tr><td>conv3-512</td></tr>
<tr><td colspan="6">Maxpool</td></tr>
<tr><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td></tr>
<tr><td rowspan="3">conv3-512</td><td rowspan="3">conv3-512</td><td rowspan="3">conv3-512</td><td>conv3-512</td><td>conv3-512</td><td>conv3-512</td></tr>
<tr><td rowspan="2">conv3-512</td><td rowspan="2">conv3-512</td><td>conv3-512</td></tr>
<tr><td>conv3-512</td></tr>
<tr><td colspan="6">Maxpool</td></tr>
<tr><td colspan="6">FC-4096</td></tr>
<tr><td colspan="6">FC-4096</td></tr>
<tr><td colspan="6">FC-1000</td></tr>
<tr><td colspan="6">soft-max</td></tr>
</table>

Figure 2: VGG19 network structure[17]

Figure 3: Improved VGG19 model

4. Work Procedure

The research flowchart is shown in (Figure 4). First, samples were scanned and digitalized into WSI images. Second, pathologists reviewed and annotated WSI. Third, the region of interest (ROI) was extracted and cut into 48×48 tumor cell images (patches), which were then inputted into models for training and verification. Fourth, the deep CNN was trained and optimized to achieve the best classification performances. Finally, the final classification results were obtained through 10-fold cross validation.

5. Datasets and Processing Methods

From the data used in this study, 35 cases of CNS-pDLBCL diagnosed by conventional pathological and clinical manifestations in the Provincial Affiliated Hospital of Shandong First Medical University In CHINA from August, 2019 to October, 2021 were chosen as the respondents. Meanwhile, 35 cases of HGG were chosen as the control group.Conventional pathological HE sections of these 70 patients were collected. Inclusion criteria were as follows: complete clinical data of patients and complete and clear paraffin pathological section. Exclusion criteria were as follows: patients whose paraffin pathological section had uneven dyeing or poor section (e.g., polluted background). The tissue sections of 70 cases were scanned using the Zhiying digital section scanning system Win20. ZYFviewer browsing program was used as the labeling tool. This program supports rectangular, circular, polygonal, and straight labels. The labeled image documents were stored in JPG format. Dataset styles are shown in (Figure 5(A)). Black polygonal labels with a line width of 2 were used. Each image was independently labeled with 1,000 tumor cells by one pathologist, and immunohistochemical CD20 dyeing results were used as reference in case of doubt. Ten cases of tumor cells were chosen randomly from 1,000 cells in each pathological section image (as is shown in Figure 5(B)). Documents were established, labeled, and stored in different classes. The original dataset was divided into training set and test set according to the ratio of 9:1.

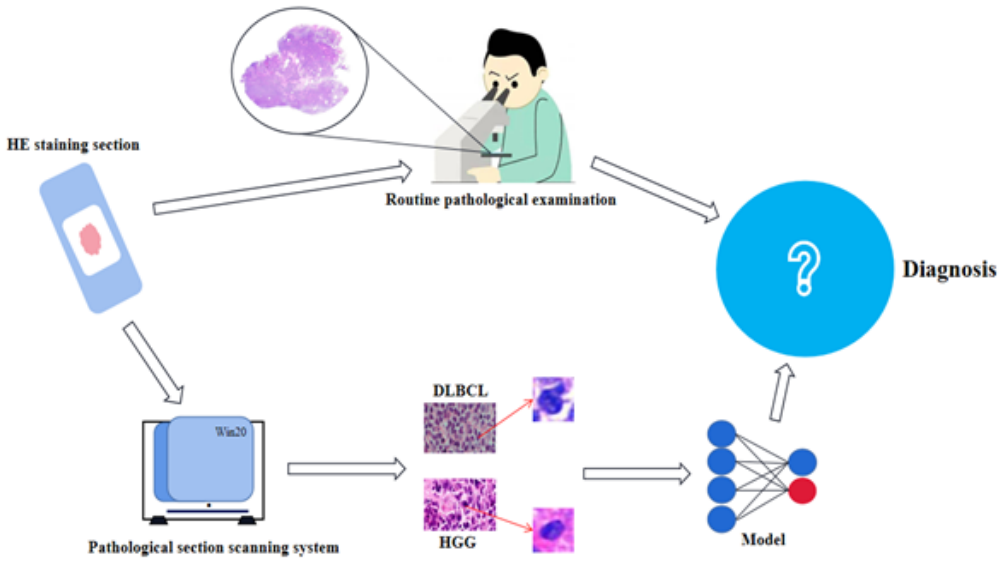

Figure 4: Work Flowchart

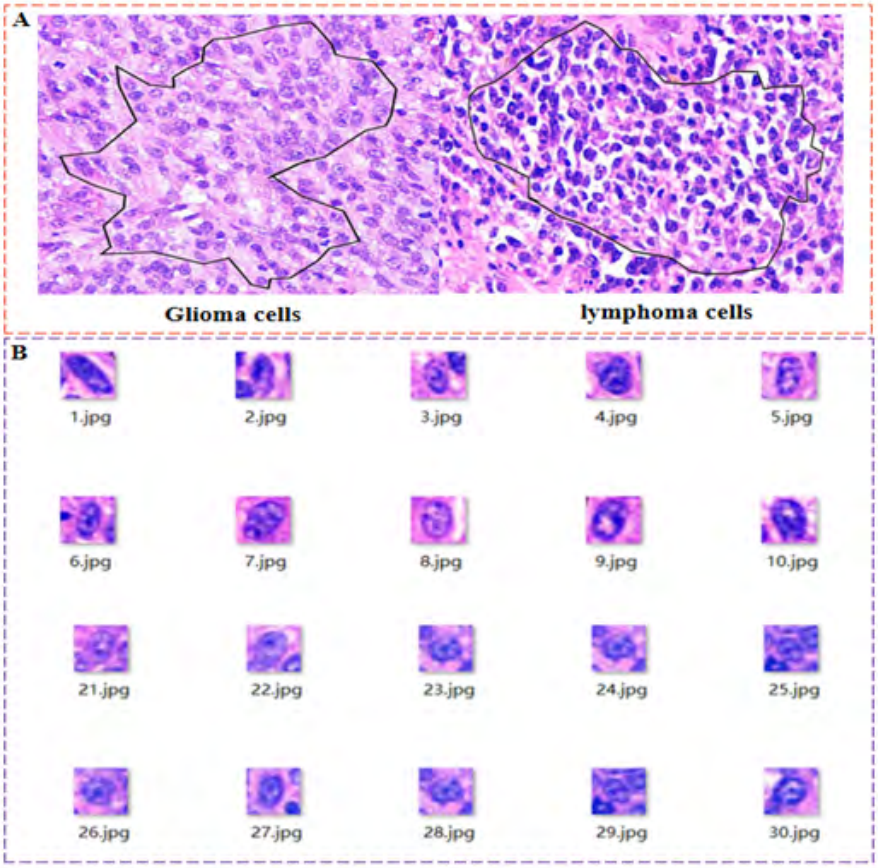

Figure 5: (A) Labeling of pathologists(The delineated area is full of tumor cells.); (B)Some images from the dataset(Cells cut from the delineated area.)

6. Model Construction

The network was built on the basis of the PyTorch deep learning framework. A class was built by using the design idea of modularization to encapsulate the whole network structure. This method can hide internal details that have been realized, increase the safety of codes, and strengthen the reuse efficiency of codes. Moreover, it is convenient to call some methods through internal integration and improve simplicity of codes. After network construction, the dataset was inputted into the network for training and classification. The results of CNS-pDLBCL and HGG cells could be obtained after a certain period.

7. Model Evaluation Standards

For the prediction results of models, manual evaluation has some subjectivity. The quantitative evaluation of the models seems impossible. The models can simply be assessed according to precision. Different evaluation methods emphasize different evaluation standards. Therefore, some relatively universal evaluation standards were chosen to evaluate the advantages and disadvantages of the models from multiple perspectives.

After training stage, the models were tested using the test dataset. Their performances were verified according to accuracy, sensitivity, specificity, FPR and F1Score. These values were determined by true positive (TP), false negative (FN), true negative (TN), and false positive (FP). The corresponding descriptions of performance indexes were introduced as follows [26-27].

$$\boldsymbol{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (1)$$

$$\boldsymbol{Sensitivity} = \boldsymbol{Recall} = \boldsymbol{TPR} = \frac{TP}{(TP + FN)} \quad (2)$$

$$\boldsymbol{Specificity} = \frac{TN}{(TN + FP)} \quad (3)$$

$$\boldsymbol{FPR} = \frac{FP}{(FP + TN)} = 1 - (3) \quad (4)$$

$$\boldsymbol{F1\ Scorce} = \frac{2TP}{(2TP + FN + FP)} \quad (5)$$

where TP is the number of positive samples predicted accurately, FP is the number of positive samples predicted wrongly, TN is the number of negative samples predicted accurately, and FN is the number of negative samples predicted wrongly. The performance of the classifier can be assessed more intuitively using the receiver operating characteristic (ROC) curve. The horizontal axis and vertical axis in the ROC space are FPR and TPR, respectively.

AUC: The magnitude of the area under the ROC curve is a way to

gauge how well a learner is doing. A unique value between 0 and 1, with larger values indicating greater performance, can be used to further assess the classifier.

7.1. Experimental Results and Analysis of this Study

We used deep convolutional network models VGG16, VGG19, Resnet18, Resnet34, and Resnet50 along with shallow classical models SVM, LeNet, and AlexNet in our experimental research to evaluate the classification performance of our model. (Figure 4) depicts the results of each model. In the experimental research, the shallow classical models SVM、LeNet、AlexNet、and deep convolutional network models VGG19、VGG16、Resnet18 、Resnet34 and Resnet50 were used to obtain preliminary classification results and compared with the new network results in order to evaluate the classification performance of our model.

8. Introduction to the Hardware

Several deep learning models were constructed on the basis of the Pytorch deep learning framework. The specific experimental environmental information is listed in the following (Table 1).

Experimental Environment	Details
Processor	Core i9 12900KF
Graphics Card	RTX3090
Deep Learning Framework	Pytorch

Table 1: Experimental environmental information

9. Results and Discussions

Each model employs a ten-fold cross-validation method to improve the classification accuracy of the data set and reduce contingency of the final result. The experimental results in (Figure 6(A)) show that the shallow model's accuracy is less than 85%, and SVM classifier's accuracy is the lowest. The reason for this may be that there are too many features and the complexity of the network model is too low. The deep network model's accuracy remains between 85% and 92% (Figure 6(B)). When the data enhancement method is not used, the performance of deep network is not outstanding and there is an obvious overfitting phenomenon. The reason for the analysis could be that the model's complexity is too high, such as the number of hidden layers being set too high, the number of neurons being set too high, and so on. When using the data augmentation method, both the shallow and deep models improved their performance, indicating that the data augmentation method can still play a similar role in increasing the amount of data in small-sample classification tasks. It is also worth noting that our model's accuracy rate can be kept between 87% and 91% without using data augmentation, while the accuracy rate has increased by roughly 6% after using data augmentation. In addition to the SVM model, the data augmentation strategy has enhanced the accuracy of other models. The average accuracy of the Resnet50 network is 91.25%, which also corresponds to the law that the shallow neural network's feature abstraction degree is low, and the deeper the level, the higher the feature abstraction degree. (Figure 6(C) shows the performance of some parameters when the accuracy rate is at its lowest, whereas (Figure 6(D) shows the performance of some parameters when the accuracy rate is at its highest. The model we use is not only more accurate than other models, but it is also somewhat superior in terms of sensitivity, specificity, and F1Score, according to the final results of the two images.

For a long time, tissue paraffin sections and immunohistochemical staining sections were the primary instruments for pathological diagnosis. Pathologists use the microscope in conjunction with their own experience to identify and diagnose lesions at the cellular level; nevertheless, the outcomes of pathological diagnosis are easily influenced by the pathologists' subjective experience, and tumor heterogeneity can occasionally be a problem. If necessary, immunohistochemical staining sections and molecular detection are required to aid diagnosis, resulting in a lengthier wait time [28]. In huge healthcare datasets, deep learning (a subset of artificial intelligence and machine learning) is being utilized to uncover patterns for disease phenotypes, event prediction, and complicated decision making [29]. At the same time that it supports pathologists in improving diagnostic accuracy, the deep learning model can minimize reporting time, labor costs, meet growing clinical needs, and tackle the problem of uneven resource allocation.

The application of deep learning in medical field can be divided into three stages: diagnosis and assisted diagnosis of tumors, precise medical treatment after tumor diagnosis, and tumor prevention by exploring evolutionary history of tumors [37]. For example, deep learning is conducive to the qualitative diagnosis of breast cancer, histological grading, lymphatic metastasis, and scoring of human epidermal growth factor receptor 2 dyeing [30-33]. Deep learning is conducive to distinguish transparent cellular renal cell carcinoma and chromophobe cell tumor [34]. In addition, a method based on automatic deep learning detection and quantitative capillaries in glioma has also been applied to accurately extract the characteristics of capillaries, thus guiding chemoradiotherapy or target treatment [35]. Nevertheless, only a few studies have been conducted on the correlation of deep learning with CNS-pDLBC and HGG. CNS-pDLBCL and HGG have poor prognosis. Early accurate diagnosis is important to the life and health of affected patients. At present, imageological examination and pathological examination are the major diagnosis techniques for CNS-pDLBCL and HGG. Imageological examination cannot determine the nature of lesions, and pathological examination is time consuming. Hence, the deep learning assisted diagnosis of CNS-pDLBCL and HGG must be urgently studied.

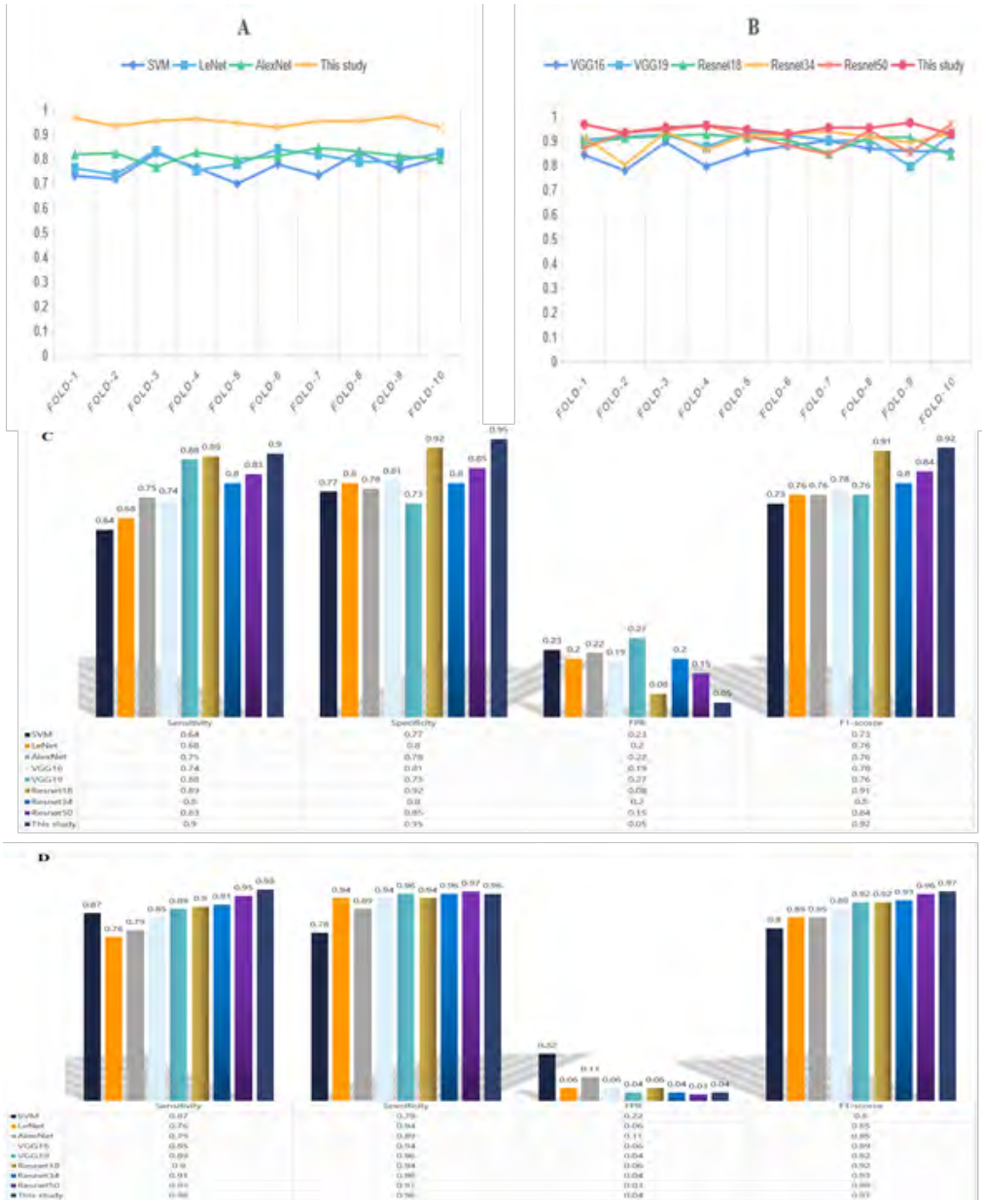

Figure 6: (A)Results of shallow model and the model in this study ; （B）Results of deep model and the model in this study ; （C）Performance indicators for models with lowest accuracy in ten-fold cross validation results ; （D）Performance indicators for models with the highest accuracy in ten-fold cross validation results.

According to our research, the model's average accuracy rate is 95%. The best strategy for improving the classification accuracy of the data set in this study is to train the data set using transfer learning and the improved VGG19. In addition, it is also verified that the use of data enhancement methods can improve the performance of all models used in this study. In 35 cases of CNS-pDLBCL and 35 cases of HGG sections, the accuracy rate of the diagnosis of pathologists who worked for 12 years was 75%, 17 cases were misdiagnosed, and in the misdiagnosed cases, there was no typical "sleeve like" structure around the blood vessels, of which 8 cases had significant cell heteromorphism and necrosis, and 9 cases had diffuse infiltration and necrosis of tumor cells in the background of nerve felt. In contrast, the model we used can obtain a variety of features that are difficult to distinguish by the naked eye from the scanning image of tumor HE section to assist in pathological diagnosis, improving the diagnostic accuracy. This shows that the ability of computer-aided diagnosis of CNS-pDLBCL and HGG tumor cells is superior to that of experienced pathologists. The findings of a study on deep learning for cytological assisted slide reading revealed that one of the causes for the model's poor judgment results was the quality of the slide itself [36]. The model missed diagnosis in the diagnosis of several tumor cells, according to the results comparison. The following factors were considered: Firstly, because the tumor was heterogeneous and the tumor cells varied, the number of samples included in this study was limited. Secondly, the image quality varies, the dyeing quality of the digital pathological section is uneven, the backdrop is contaminated, or the presence of artifacts in the section impairs the model's performance.

10. Conclusion

In this study, we created and tested a model for the differential diagnosis of CNS-pDLBCL and HGG pathological pictures. This model can help pathologists make objective, accurate, and timely diagnoses of CNS-pDLBCL and HGG, as well as providing a novel research idea for future CNS-pDLBCL and HGG pathology. At the same time, our research shows that using transfer learning and data augmentation methods can reduce over-fitting, improve training speed and accuracy, and raise the model's generalization capacity. We plan to expand and improve our research on the classification of CNS-pDLBCL and HGG in the future. Begin with the following aspects: First, the standardization and unification of the manufacturing process, such as material collection, fixation, and dyeing;.Second, the input, storage, and transmission of high-quality slices. Third, the input, storage, and transmission of high-quality slices, and then create a big and high-quality digital pathological segment database. Fourth, by merging imaging and biological indicators, this model's diagnosis accuracy can be improved further.

11. Limitations

However, we don't know whether our neural network learning what to write, we don't know what is the meaning of each parameter, so we can't explain the entire model of operation mechanism, therefore, we need to find some better ways to make up for the defect.

12. Funding

This work was supported by the National Key R&D Program of China(2019YFE0117800), the National Natural Science Foundation of China(82071035), the National Natural Science Foundation of China(22176115),Shandong Provincial Key Research and Development Program (Major Scientific and Technological Innovation Project) (2021CXGC010506), Shandong Provincial Natural Science Foundation (ZR2021MH208), Shandong Provincial Natural Science Foundation (ZR2021QH290).

13. Acknowledgement

The authors would like to thank all the reviewers who participated in the review.

14. Conflicts of Interest

The authors declare that they have no conflicts of interest.

15. Consent to Publish

All authors consent to publish this article.

16. Authors Contribution

All authors have participated in this study, and consent to publish this article in Journal. The contribution list is shown in the following form. Guarantor of integrity of entire study: Miaoqing Zhao and Wei Li. Study concept: Miaoqing Zhao and Wei Li. Study design: Wei Li and Miaoqing Zhao. Literature research: Yue Xi, Chongxuan Tian Cong Wu, Yuting Ma. Data acquisition: Yue Xi. Data analysis: Chongxuan Tian. Manuscript preparation: Chongxuan Tian and Yue Xi. Manuscript editing：Cai Chen and Kun Ru. Manuscript revision/review: Chongxuan Tian.

17. Ethical Statement

This study was approved by the ethics committee of Shandong Provincial Hospital affiliated to Shandong First Medical University. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

1. Hoangxuan K，Bessell E，Bromberg J. Diagnosis and treatment of primary CNS lymphoma in immunocompetent patients: guidelines from the europeanassociation for neuro oncology. Lancet Oncol. 2015; 16(7) : 322- 332.
2. Yao L, Jie P, Liu Y. The value of 3.0T magnetic resonance diffusion tensor imaging in the differential diagnosis of high-grade glioma and brain metastases. Journal of Clinical Radiology. 2020; 39(1): 22-25.
3. Suh CH，Kim HS，Jung SC. MRI as a diagnostic biomarker for differentiating primary central nervous system lymphoma from glioblastoma: a systematic review and meta-analysis. J Magn Reson Imaging. 2019; 50: 560-572.
4. Du E, Xu L, Zhou X. Discussion on the diagnostic value of plain CT and enhanced MRI in primary central nervous system lymphoma. Practical Hospital Clinical Journal. 2018; 15: 246-247.
5. Yang Z, Leng S, Duan W. Clinical significance of MRI in the differential diagnosis of primary central nervous system lymphoma and multiple brain glioma. Journal of Clinical and Experimental Medicine. 2019; 18(5): 548-551.
6. Zhang Z, Wang J, Qu H. To explore the value of MRI in the differential diagnosis of primary central nervous system lymphoma and glioma. Yingxiang Yanjiu Yu Yixue Yingyong. 2021; 5(03): 173-174.
7. Xu Shaohua. CT and MRI manifestations and clinicopathological features of primary central nervous system lymphoma. Shenzhen Journal of Integrated Traditional Chinese and Western Medicine. 2021; 31(18): 110-112.
8. Liu R, Shi K, Wang L. Analysis of differential diagnosis of primary central nervous system lymphoma and brain glioma by CT and MRI. Chinese Journal of Clinical Oncology and Rehabilitation. 2022; 29(04): 385-389.
9. Geng L, Sun Y, Sun Z, Wang X, Xu K. MRI differential diagnosis of primary central nervous system lymphoma and high-grade glioma with deep lesions. Chinese Journal of Interventional Imaging and Therapy. 2018; 15(02): 95-99.
10. Suh CH，Kim HS，Jung SC. MRI as a diagnostic biomarker for differentiating primary central nervous system lymphoma from glioblastoma: a systematic review and meta-analysis. J Magn Reson Imaging. 2019; 50: 560-572.
11. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;

521(7553): 436–444.

12. Deng L, Yu D. Deep learning: methods and applications. Found Trends Signal Process. 2014; 7: 197-387.
13. Wang S, Yang DM, Rong R. Artificial intelligence in lung cancer pathology image analysis. Cancers (Basel). 2019; 11: E1673.
14. Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J, et al. AI-based pathology predicts origins for cancers of unknown primary. Nature. 2021; 106-110.
15. Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems. 2012; 25(2).
16. Lecun Y, Bottou L. Gradient-based learning applied to document recognition. Proceedings of the IEEE. 1998; 86(11): 2278-2324.
17. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv. 2015; 1409.1556.
18. Noble WS. What is a support vector machine?. Nature Biotechnology. 2006; 24(12): 1565-1567.
19. Li Y. A review on the application of transfer learning in medical image analysis. Computer Engineering and Applications. 2021; 57(20): 42-52.
20. Kim HE, Cosa-Linan A, Santhanam N. Transfer learning for medical image classification: a literature review. BMC Medical Imaging. 2022; 22(1): 1-13.
21. Perez L, Wang J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. 2017.
22. Chlap P, Min H , Vandenberg N. A review of medical image data augmentation techniques for deep learning applications. Journal of Medical Imaging and Radiation Oncology. 2021.
23. He K , Zhang X, Ren S. Deep Residual Learning for Image Recognition. IEEE. 2016.
24. Tang Y . Deep Learning using Linear Support Vector Machines. Computerence. 2013.
25. Xiong F, He Di, Liu Y. Classification of Pneumonia Images Based on Improved VGG19 Convolutional Neural Network (Invited. Acta Photonica Sinica. 2021; 50(10): 1010001
26. Geng L, Zhang S, Tong J, Xiao Z. Lung segmentation method with dilated convolution based on VGG-16 network. Comput Assist Surg (Abingdon). 2019; 24(sup2): 27-33.
27. Zadeh Shirazi A, Fornaciari E, Bagherian NS, Ebert LM, Koszyca B, Gomez GA, et al. DeepSurvNet: deep survival convolutional network for brain cancer survival rate classification based on histopathological images. Med Biol Eng Comput. 2020; 58(5): 1031-1045.
28. Li M , Ma X , Chen C. Research on the Auxiliary Classification and Diagnosis of Lung Cancer Subtypes Based on Histopathological Images. IEEE Access. 2021; PP(99): 1-1.
29. Ninon B, Simona B, Johann F. Deep learning for brain disorders: from data processing to disease treatment. Briefings in Bioinformatics. 2020.
30. Jia M, Guo X, Tian F. Application of Artificial Intelligence Technology in Pathological Image Analysis of Breast Tissue. Journal of Physics: Conference Series. IOP Publishing. 2020; 1642(1): 012018.
31. Cruz-Roa A, Gilmore H, Basavanhally A. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. Sci Rep. 2017; 7: 46450.
32. Liu Y, Kohlberger T, Norouzi M. Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists. Arch Pathol Lab Med. 2019; 143(7): 859-868.
33. Spanhol FA, Oliveira LS, Petitjean C. A Dataset for Breast Cancer Histopathological Image Classification. IEEE Trans Biomed Eng. 2016; 63(7): 1455-1462.
34. Tabibu S, Vinod PK, Jawahar CV. Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Sci Rep. 2019; 9(1): 10509.
35. Chang HY, Jung CK, Woo JI. Artificial intelligence in pathology. Journal of pathology and translational medicine. 2019; 53(1): 1-12.
36. Hu A, Zhu L, He H. The application value of artificial intelligence-assisted analysis technology in the screening of cervical precancerous lesions. Chinese Journal of Clinical and Experimental Pathology. 2022; (01): 27-30.
37. Yu G, Wei P, Chen Y. Artificial intelligence in pathological diagnosis and assessment of human solid tumor: Application and thinking. Acad J Sec Mil Med Univ. 2017; 38: 1349-1354.