Postprint: WSI-Based Classification of Breast Pathology Subtypes

Authors: Chen Jinling, Li Jie, Zhao Chengming, Liu Xin

Date: 2022-05-11T10:48:43Z

Abstract

To achieve accurate classification of breast pathology WSI images, we propose a gated convolutional neural network classification method based on hybrid connections. A hybrid module incorporating local residual connections and global dense connections is constructed, with squeeze-and-excitation gated units embedded within it, establishing a backbone network where hybrid modules and transition layers are connected alternately. The model is trained using an image data augmentation method based on quadtree segmentation. Experimental results on the BreastSet clinical dataset demonstrate that the proposed method achieves image-level, patient-level, and pathology-level accuracies of 92.24%, 92.83%, and 92.18%, respectively. Compared with other methods, this approach exhibits improved accuracy while reducing the number of parameters and computational cost, thereby offering greater clinical application value.

Full Text

Preamble

Research of Breast Pathological Subtype Classification on WSI

Chen Jinling†, Li Jie, Zhao Chengming, Liu Xin
(School of Electrical Information, Southwest Petroleum University, Chengdu 610599, China)

Abstract: To achieve precise classification of pathological breast Whole Slide Image (WSI) images, this study proposes a gated convolutional neural network classification method based on hybrid connections. The approach constructs a hybrid module that combines local residual connections with global dense connections, embeds a squeeze-excitation-gated unit into the hybrid module, and establishes a backbone network with alternating hybrid modules and transition layers. Combined with an image data augmentation method based on

quad-tree segmentation for model training, experimental results on the Breast-Set clinical dataset demonstrate that the proposed method achieves image-level, patient-level, and pathology-level accuracies of 92.24%, 92.83%, and 92.18%, respectively. Compared with alternative methods, this approach improves accuracy while reducing parameter count and computational requirements, offering greater clinical application value.

Keywords: WSI; breast pathological subtype classification; computer-aided diagnosis; gated convolutional network; hybrid connection

---

0 Introduction

Breast cancer ranks as the most prevalent malignant tumor among Chinese women, with incidence rates rising annually and posing a severe threat to women's health and lives. Early-stage breast cancer is a curable disease, and accurate diagnosis can maximize patient survival probability and quality of life. With the development of intelligent algorithms, accumulation of medical data, and advancement of healthcare capabilities, various intelligent algorithms have gradually been applied in the medical field. Traditional image classification algorithms consist of three steps: feature extraction, feature encoding, and classifier design. The manual extraction of pathological features typically consumes substantial computational resources and is difficult to implement practically. Convolutional Neural Networks (CNN) possess excellent automatic feature modeling capabilities. Currently, CNNs are widely used to construct deep learning (DL) models for breast pathological image classification.

Histopathological images of breast tissue play a significant role in clinical diagnosis. The Breast Cancer Histopathological Database (BreakHis) is a large public dataset that has already been partitioned, and numerous researchers have conducted pathological classification studies based on it. For instance, Han et al. constructed a CSDCNN model using deep learning model foundation layers, achieving an average classification accuracy of 93.3% in multi-classification tasks. Bardou et al. employed manual feature extraction methods based on Bag of Words, Locality Constrained Linear Coding, and Support Vector Machine (SVM) classifiers, obtaining accuracies of 96.15% and 93.31% for binary and multi-classification tasks, respectively. Their second approach, a deep learning classification method based on data augmentation and convolutional neural networks, achieved accuracies of 98.33% and 88.23% for binary and multi-classification tasks, respectively. Cheng et al. utilized two deep learning models, VGG16 and InceptionV3, combined with transfer learning and data augmentation methods, achieving accuracies of 94.43% and 82.64% for benign-malignant binary classification, and 82.97% and 58.78% for eight-class pathological subtype classification. Nguyen et al. established a CNN model incorporating resized original images and other methods for automatic classification of multiple breast cancer types, achieving an accuracy of 73.68%. Mt et al. processed each image

dataset using image enhancement techniques, selected and processed important key regions of images through attention modules, and established the BreastNet model, which achieved a maximum accuracy of 98.80% for benign-malignant binary classification. Li Yuqian used deep learning methods as a feature extractor and constructed a random forest model based on breast image features, achieving a multi-classification accuracy of 92.5%. Jiang et al. established a convolutional neural network model based on small SE-ResNet modules, achieving accuracies of 98.87%~99.34% and 90.66%~93.81% for benign-malignant binary classification and eight-class pathological subtype classification, respectively. Liu Qiaoli et al. proposed a breast cancer classification model based on DenseNet, achieving an average recognition accuracy of 99.2% in binary classification tasks. Liu Jingwen et al. proposed a breast cancer pathological image recognition method based on Inception-ResNet-V2, achieving an eight-classification accuracy of 79.7%. Ming Tao et al. combined attention mechanism algorithms and proposed a deep learning-based multi-scale channel recalibration model, which demonstrated an accuracy of 88.87% in benign-malignant binary classification experiments. Yu Lingtao et al. proposed a multi-task model based on convolutional neural networks, achieving classification accuracies of 98.55%~99.52% in binary tasks and 92.26%~94.85% in multi-classification tasks. Zhao Xiaoping et al. proposed a multi-classification model for breast cancer histopathological images based on dense convolutional neural networks, attention mechanisms, and focal loss functions, achieving a benign-malignant binary classification accuracy of 99.1% and an eight-class pathological subtype classification accuracy of 95.5% on the BreakHis dataset.

In clinical diagnosis of breast pathological types, pathologists typically observe and analyze breast pathological WSI (Whole Slide Image) images directly. Therefore, research on breast pathological WSI image data holds greater practical value. Cruz-Roa et al. proposed a deep learning method for automatic detection and visual analysis of invasive ductal carcinoma tissue regions in whole-slide breast pathological images, achieving a balanced accuracy of 84.23% on 49 WSI images. Wang et al. designed a metastatic breast cancer detector for automatic detection of sentinel lymph node WSI images, achieving an image classification AUC of 0.995 and a tumor localization score of 0.733. Gecer et al. established a deep learning model composed of four fully convolutional networks for five-class classification of breast WSI images, with prediction results showing 55% similarity to diagnoses by 45 pathologists.

To achieve accurate and efficient classification of pathological subtypes in breast whole-slide images, this paper utilizes breast pathological WSI images from Mianyang Central Hospital to establish the BreastSet dataset and proposes a gated convolutional network model based on hybrid connections (HC-GCN) for classification research. The model employs a hybrid connection architecture that accelerates forward information propagation through cross-layer "shortcut" connections, reducing the massive redundancy caused by dense connections. It utilizes a gating structure combining attention mechanisms to effectively fuse feature information, thereby improving pathological category discrimination ac-

curacy and reducing computational complexity of the classification model. Additionally, this paper employs the Tanh' function as the activation function for the hybrid module characteristics, avoiding the issue of neuron "death" caused by the ReLU function that prevents further training. The model performance is enhanced from both network structure and activation function perspectives.

1.1 Hybrid Connections

The residual connections proposed by He et al. broke through the development bottleneck of deep learning, enabling the training of effective deep neural networks through cross-layer data channel structures and establishing the ResNet model. However, in practice, ResNet typically involves numerous network layers, so despite its internal parameter sharing mechanism, the parameter count and computational load remain substantial. Additionally, during the stacking of residual modules, early feature information is naturally lost, and excessive residual aggregation leads to excessive feature information loss, ultimately resulting in irreparable errors. DenseNet, proposed by Huang G et al., is a network model based on dense connections. Like ResNet, it employs cross-layer connection patterns, but differs in that the output of residual blocks is the sum of output and input, while the output of dense blocks is the concatenation of output and input along the channel dimension. DenseNet alleviates the gradient vanishing problem, promotes effective transmission and utilization of feature information, preserves more feature information, requires less network depth, and features significantly fewer channels per network layer, reducing model parameters. However, in dense connection networks, each layer' s feature information aggregates all previous layers. While this approach protects feature information utilization, it also causes the model to repeatedly extract redundant information, resulting in extremely high redundancy and persistent computational load.

To improve breast cancer detection accuracy while reducing algorithm complexity and computation time, this paper proposes a hybrid connection that combines local residual connections with global dense connections. The basic structure of the hybrid connection module is illustrated in Figure 1.

Figure 1 shows a hybrid module composed of two SEG blocks connected in a hybrid manner, with each hybrid module consisting of 2 or more SEG blocks. Solid arcs represent dense connections, dashed arcs represent residual connections, "+" denotes the addition of two inputs, "[,]" denotes concatenation along the channel dimension, and and represent the compression unit and activation unit in the SEG block, respectively.

The ultimate goal of the hybrid module is to replace the original dense block in the DenseNet network, implementing hybrid connections in the network model and enhancing the feature extraction capability of the pathological classification model for breast pathological images.

1.2 Squeeze-Excitation-and-Gated Unit

SEG1 and SEG2 in Figure 1 represent special structures built specifically for hybrid modules, similar in function and principle to bottleneck layers, called SEG blocks. The specific structure of the SEG block is shown in Figure 2.

Let H, W, and C denote the height, width, and number of channels of the feature map, respectively. As shown in Figure 2, in the SEG module, convolutional and group convolutional structures are first used to compress the input feature map, controlling feature dimensions, reducing model channel numbers, and thereby decreasing model parameters and computational requirements.

Secondly, forget gate and update gate structures incorporating attention mechanisms are employed to effectively process feature information. The forget gate uses a global context model to calculate spatial attention feature weights for each position and adds a Sigmoid function for weighted attenuation of each channel, effectively reducing computational requirements. The update gate uses the Softmax function for comparative selection, identifying the most salient feature information to overlay with original information, promoting model exploration and application of new features, and improving classification accuracy to a certain extent. Simultaneously, the structures of both update and forget gates can effectively model global context like self-attention mechanism models while conserving computational resources. The following sections introduce the three structures in the SEG block: compression and activation, forget gate, and update gate.

1.2.1 Compression and Activation

The compression unit operation consists of two main steps: First, a 1×1 standard convolution is used to compress the input feature map to reduce channel numbers, decreasing from C to αC, where α represents the channel width. Then, a 3×3 group convolution with group number g further compresses the feature map, with stride settings enabling downsampling operations to reduce feature map height, width, and channel numbers. The fundamental purpose of both compression operations is to reduce model parameters and computational load.

After feature information is compressed, it enters the activation unit, where 3×3 depthwise convolution is employed for activation and data padding. Hidden information in breast pathological images is substantially more abundant than in common classification images (e.g., cats, dogs, flowers). Extracting more information requires more convolutional kernels in convolutional layers, which increases computational scale. Under the same input-output feature map dimensions, depthwise convolution requires very few input channels and has extremely low parameter and computational costs, with significantly fewer parameters and computations compared to standard convolution. Using depthwise convolution for activation and data padding can enhance model performance without affecting efficiency.

1.2.2 Forget Gate

The leftmost connection line in Figure 2 and the dashed line in Figure 1 both represent residual connections within the SEG block. Unlike ordinary residual connections that simply add channel information to residual connection information, this implementation simultaneously uses both residual and dense connections. To efficiently utilize repeatedly transmitted feature information in channels, a forget gate incorporating attention mechanisms is embedded in the residual connection to attenuate and filter reused feature information.

To satisfy the attenuation requirement for reused features, the forget gate must ensure effective feature information flow while mapping each channel weight's final output within the range (0, 1). Therefore, the forget gate employs both spatial and channel attention mechanisms while using the Sigmoid function for final attenuation.

The forget gate structure is shown in Figure 3, with operation steps divided into five main phases:

Step 1: The feature map X is input into a 1×1 standard convolution, outputting a spatial attention feature S' with reduced channel dimensions.

Step 2: The Softmax function normalizes the spatial attention feature map to obtain a new spatial attention feature S, where each element can be represented by the formula:

$$S_{i,j} = \frac{e^{S'_{i,j}}}{\sum_{i=1}^{H} \sum_{j=1}^{W} e^{S'_{i,j}}}$$

Step 3: The feature map X passes through a global attention pooling layer, generating a reduced global feature map z. The representation formula for the global feature map of channel c is:

$$z_c = \sum_{i=1}^{H} \sum_{j=1}^{W} S_{i,j} \cdot X_{i,j,c}$$

Step 4: The global attention feature map z passes through two consecutive fully connected layers. Since parameter updates during network training cause input data distribution changes in subsequent layers, a Batch Normalization + Tanh activation layer is added between the two fully connected layers for buffering, performing batch normalization operations between the layers.

Step 5: The Sigmoid function performs final attenuation. The final output feature map of the forget gate is represented by:

$$f = \sigma(W_{f,2} \cdot \text{Tanh}(\text{BN}(W_{f,1} \cdot z + b_{f,1})) + b_{f,2})$$

where f represents the final output feature map of the forget gate, $W_{f,i}$ and $b_{f,i}$ represent the weights and bias values of the i-th fully connected layer in the

forget gate, σ represents the Sigmoid function, and r_f represents the bottleneck ratio number set for the forget gate, with C/r_f denoting the number of neurons in the intermediate layer.

1.2.3 Update Gate

The update gate is designed to more effectively process new features obtained from 3×3 depthwise convolution. The specific structure is shown in Figure 4. The difference between the update gate and forget gate lies in their purposes: the forget gate aims to attenuate information, while the update gate aims to promote feature information mining and utilization. Structurally, the update gate does not use the Sigmoid function for feature attenuation and includes a simple addition operation between its final output and the global attention pooling layer output.

Let h represent the hidden feature map output after the Batch Normalization + activation function layer, where $h \in \mathbb{R}^{C/r_u \times 1 \times 1}$, and r_u represents the bottleneck ratio number set for the update gate. The specific formula for the hidden feature map is:

$$h = \text{Tanh}(\text{BN}(W_{u,1} \cdot z + b_{u,1}))$$

where $W_{u,1}$ and $b_{u,1}$ represent the weights and bias values of the first fully connected layer, respectively.

The hidden feature map h generates a channel attention feature map v after passing through the second fully connected layer:

$$v = W_{u,2} \cdot h + b_{u,2}$$

where $W_{u,2}$ and $b_{u,2}$ represent the weights and bias values of the second fully connected layer, respectively.

The output v of the second fully connected layer is added to the output z of the global attention pooling layer to obtain the final output X' of the update gate:

$$X' = v + z$$

Essentially, the update gate and forget gate treat the compact feature map z formed by the compression unit as reused old features and the output X'' formed by 3×3 depthwise convolution as extracted new features, finally extracting and aggregating old and new features to constitute the final output X'. This processing method not only promotes effective feature reuse while saving parameters and computations but also enhances the ability to mine more feature information.

1.2.4 Activation Function Selection

Activation functions introduce nonlinear factors into neural network training processes, improving model feature expression capabilities. Figure 5 shows the

function images of three commonly used activation functions: ReLU, Sigmoid, and Tanh.

The ReLU function offers advantages such as alignment with neural network coding characteristics, overcoming gradient descent problems, and improving computational efficiency. However, the ReLU function curve in Figure 5 reveals two critical drawbacks. First, when input neurons $x < 0$, ReLU's gradient is zeroed out, which can likely cause neuron "death" where input data cannot be activated. Second, ReLU function outputs are always positive without negative values, significantly undermining the model's feature representation capability and preventing effective feature information training. Additionally, ReLU functions do not control numerical magnitude, and excessive magnitude variation in deep networks may prevent model training.

Sigmoid and Tanh functions effectively solve the neuron "death" problem, protect feature information, and control numerical magnitude. However, as shown in Figure 5, when using Sigmoid as the activation function, if the current input parameter $x < 0$, the optimal optimization direction is $(\omega_1, \omega_2) + (-d, +d)$. Since the Sigmoid function is not zero-centered and its output values are always positive, the model cannot achieve the fastest parameter update but instead approaches the optimal solution in a zigzag pattern. Therefore, using the Tanh function, which is zero-centered with a value range of $[-1, 1]$, can accelerate model convergence. Moreover, Tanh's gradient vanishing problem is less severe than Sigmoid's, but Tanh's approximate linearity in the $[-1, 1]$ interval may cause classification errors and reduce feature learning accuracy.

To address Tanh's gradient vanishing and classification error issues, this paper selects Tanh' as the activation function for HC-GCN. The specific formula is:

$$\text{Tanh}'(x) = \text{Tanh}(x) \cdot \text{sigmoid}(x)$$

Using the Tanh' function as the activation function can effectively solve the "neuron death" problem, protect feature information, control numerical magnitude, eliminate classification errors caused by approximate linear processing, and achieve better model performance and classification results. Additionally, hybrid connections significantly strengthen gradient backpropagation, largely compensating for potential gradient vanishing issues caused by activation functions.

1.3 Overall Model Architecture

The overall architecture of HC-GCN is similar to DenseNet, with the main difference being that DenseNet's primary building block is the dense block, while HC-GCN's primary building block is the hybrid block. Both hybrid blocks and dense blocks define the input-output connection patterns of bottleneck blocks within the model.

The HC-GCN model takes breast pathological images requiring pathological category classification as network input. First, input images undergo standardization processing through a feature block composed of a 3×3 convolutional layer, Batch Normalization (BN), and ReLU function. Then, features are passed into a backbone network where hybrid modules and transition layers are interconnected, with transition layers primarily controlling channel numbers. Finally, extracted feature information enters a Global Average Pooling (GAP) layer, fully connected layer, and Softmax classifier to complete classification. Figure 6 shows the HC-GCN backbone network composed of 3 hybrid modules and 2 transition layers in alternating connection, where each hybrid module internally contains multiple Squeeze-Excitation-and-Gated modules (SEG) connected in dense mode.

In practical applications, to obtain better pathological category classification results for breast pathological images, different HC-GCN structures can be constructed by varying the number of hybrid modules, SEG blocks, and adjusting hyperparameters.

2.1 Dataset Sources

The image data used in this paper are breast pathological WSI images provided by Mianyang Central Hospital, with source files in KFB format that can be viewed and processed using K-Viewer software. Figure 7 shows partial samples from the BreastSet dataset.

After exporting original data as RGB three-channel images in PNG format, a classification dataset is formed containing 3,498 WSI pathological images from patients with different breast cancer conditions. Classification labels include eight categories: Medullary Breast Carcinoma (MBC), Non-Special Type Invasive Breast Carcinoma (NST), Apocrine Carcinoma (AC), Invasive Micropapillary Carcinoma (IMPC), Invasive Lobular Carcinoma (ILC), Mucoid Carcinoma (MC), Solid Papillary Carcinoma (SPC), and Tubular Carcinoma (TC). This dataset is named the Breast Pathological Image Classification Dataset (Breast-Set). Table 1 presents the data structure of the BreastSet dataset.

2.2 Data Augmentation

WSI images have large dimensions and memory footprints, making image data in the BreastSet dataset difficult for computers to directly process and apply. To ensure accurate image data reading and improve model feature expression capability, data augmentation can be applied to the dataset. In addition to basic methods such as resizing, the quad-tree segmentation-based image data augmentation method shown in Figure 8 can also be used.

Quad-tree is a tree data structure where each node has four sub-blocks, commonly applied in two-dimensional spatial data analysis and classification. It divides data into four quadrants, with data ranges that can be square, rectan-

gular, or any other shape. Quad-tree has a continuous structure where, at each level, the input image from the previous level is equally divided into four parts. After segmentation, each sub-image is considered to have the same class label as the original image.

2.3 Fusion Algorithm

Due to the quad-tree image segmentation method, one image is divided into multiple sub-image blocks, each of which may produce different classification results after model computation. Therefore, a fusion algorithm is needed to effectively integrate classification results from each image block. Commonly used fusion algorithms include sum rule, product rule, max rule, and majority voting rule. The specific calculation formulas for sum rule, product rule, and max rule algorithms are as follows:

Sum Rule:

$$\phi = \arg\max_k \sum_{i=1}^{N} p_{ik}$$

Product Rule:

$$\phi = \arg\max_k \prod_{i=1}^{N} p_{ik}$$

Max Rule:

$$\phi = \arg\max_k \max_i p_{ik}$$

where p_{ik} represents the probability value that the i-th sub-image block of one image is classified as class k by the model; K represents the total number of pathological categories; and N represents the total number of split blocks for one image.

This paper employs an integration method combining multiple fusion algorithms. Figure 9 shows the complete fusion classification result process for one case image, where the final output classification result is determined by voting among classification results obtained from the three fusion rules.

2.4 Evaluation Metrics

Recording model training and testing results provides the basis for further model improvement and quality evaluation. To assess model applicability for breast pathological subtype classification tasks on WSI, this paper calculates classification accuracy, model parameters (params), floating-point operations (FLOPs), precision (P), recall (R), and F1 scores for evaluation.

In the medical field, classification accuracy of computer-aided diagnosis models includes three types: image-level, patient-level, and pathology-level. Image-level accuracy (IA) refers to the ratio of correctly classified images to total sample

images. Patient-level accuracy (PA) refers to the average classification accuracy of breast pathological images corresponding to each patient. Pathology-level accuracy (PLA) refers to the average classification accuracy of breast pathological images for each pathological type in multi-classification scenarios.

Precision, recall, and F1 score are important metrics for measuring model performance. Precision is the probability that patients diagnosed as positive are true positives, while recall is the probability that true positive patients are diagnosed as positive. High precision reduces misdiagnosis rates, preventing healthy individuals from receiving unnecessary treatment and wasting medical resources. High recall reduces missed diagnosis rates, preventing patients from missing optimal treatment windows.

The F1 score is an important metric in statistics for measuring model accuracy and represents a harmonic mean of precision and recall. Specific calculation formulas are shown in equations (5)~(7):

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

where TP represents true positive count, FP represents false positive count, and FN represents false negative count.

In multi-classification pathological scenarios, macro-average and micro-average values must be calculated for precision, recall, and F1 scores. Macro-average refers to calculating the average of precision, recall, and F1 scores for each pathological category separately, treating each category equally and thus being susceptible to categories with fewer samples. Micro-average refers to calculating precision, recall, and F1 scores overall without distinguishing categories, treating each sample equally and thus being susceptible to categories with more samples. By calculation principle, Micro-P, Micro-R, and Micro-F1 are numerically equal to IA.

3.2 Model Architecture and Hyperparameter Settings

To achieve better model classification performance, after multiple rounds of testing and adjusting the number of SEG blocks in hybrid modules and growth rate parameters, an HC-GCN model was constructed with 3 hybrid modules, each containing 8 SEG blocks. Table 3 shows the specific structural settings of the model.

Additionally, several important hyperparameters are set as follows: batch size is 16, epoch is 100, learning rate is 0.1, compression factor θ between transition

layers and hybrid modules is 0.5; group number g for 3×3 group convolution in SEG blocks is 4, width multiplier α is 4, reduction ratios for update gate and forget gate (r_u and r_f) are both 2, stride S is 1; in transition layers, compression factor θ is set to 1, expansion factor α is 1.5, stride S is 2; depthwise convolution kernel size is 2.

3.3 Training and Testing

As shown in Table 1, before experiments, the training set, validation set, and test set were allocated in an 8:1:1 ratio. This paper trained on level 0 and level 1 of breast pathological image data augmentation, meaning training was conducted in two modes: directly using original data and training after quad-splitting images, where each original image was cut into 4 blocks for training and voting fusion.

Figure 10 shows the training set loss curve and validation set image-level accuracy curve obtained by HC-GCN based on the BreastSet dataset. The upward-curving solid line represents the validation set image-level accuracy curve for original data classification (Acc0), the downward-curving solid line represents the training set loss curve for original data classification (Loss0), the upward-curving dotted line represents the validation set image-level accuracy curve for quad-split training mode (Acc1), and the downward-curving dotted line represents the training set loss curve for quad-split training mode (Loss1).

Figure 10 demonstrates that HC-GCN achieves good accuracy results and low loss when trained on both original and quad-split datasets, indicating that the established model possesses certain accuracy and scientific validity. The model's training and validation loss curves show a relatively stable training process on the BreastSet dataset, demonstrating model stability. When epoch is greater than or equal to 40, accuracy and loss values become relatively high and gradually converge, indicating model convergence.

Comparing the validation image-level accuracy curves of the two training modes in Figure 10, both training modes show smooth training accuracy curves with high accuracy values, demonstrating that the HC-GCN model is suitable for subtype classification research on the current breast pathological dataset.

3.4 Experimental Results Analysis

Table 4 shows the classification experimental results of the proposed method and other classical deep learning methods based on the BreastSet dataset.

From the data corresponding to various models in original data training mode and quad-split training mode, all accuracy performance metrics including accuracy and F1 scores obtained in quad-split training mode are superior to those in original data training mode. Simultaneously, Figure 10 shows that compared with original data training mode, quad-split training mode generally yields higher validation set image accuracy curves and lower training set loss

curves, proving that the quad-tree segmentation-based image data augmentation method helps improve model classification performance for breast pathological WSI images.

HC-GCN achieves good pathological type recognition accuracy using very few parameters and computations, helping reduce model dependence on computational resources and increasing model scalability and portability.

3.5 Comparison with Other Methods

This paper also selected multiple classical and popular image classification models for comparison with HC-GCN. Each model performed pathological subtype classification based on the BreastSet dataset, with experimental results listed in Table 4.

Table 4 shows that many excellent models can achieve high classification accuracy on breast pathological image classification tasks. Compared with DenseNet121, which has the highest accuracy at each level, HC-GCN improves image-level, patient-level, and pathology-level accuracies by 1.14%, 0.23%, and 0.65%, respectively, in original data training mode, and by 0.37%, 0.40%, and 0.65%, respectively, in quad-split training mode. This demonstrates that HC-GCN can achieve higher classification accuracy than the DenseNet model while requiring only approximately 3/8 of the parameters and 1/6 of the floating-point operations of DenseNet121, effectively alleviating the excessive redundancy problem caused by feature reuse mechanisms and saving computational resources.

Compared with ShuffleNetv2, which has a smaller parameter count, HC-GCN achieves improvements of 9.77%, 9.53%, and 9.45% in image-level, patient-level, and pathology-level accuracies, respectively, in original data training mode, and improvements of 5.57%, 3.53%, and 17.61%, respectively, in quad-split training mode, while using even fewer parameters and computations.

Compared with AlexNet, MobileNetv2, and MobileNetv3, which have smaller computational requirements, although HC-GCN uses 0.19GB~0.28GB more computations, it improves image-level, patient-level, and pathology-level accuracies by 5.46%~19.54%, 5.84%~16.89%, and 6.66%~29.32%, respectively, in original data training mode, and by 6.24%~12.33%, 5.93%~12.37%, and 5.24%~11.75%, respectively, in quad-split training mode.

Compared with other models listed in Table 4, HC-GCN achieves superior accuracy, precision, recall, and F1 score values while using fewer parameters and computations. Experimental results demonstrate that HC-GCN significantly improves subtype classification performance for WSI breast pathological images.

In summary, compared with other methods, HC-GCN not only improves pathological category classification accuracy for breast pathological images but also reduces required model parameters and computations, saving computational resources.

3.6 Ablation Experiments

To test and verify the effectiveness of the proposed methods and modules, four sets of comparative ablation experiments were designed. Table 5 shows the sequential addition of improved structural modules in the ablation experiments.

In Table 5, RC represents residual connection, FG represents forget gate, and UG represents update gate. In the first ablation experiment, each bottleneck block contains one standard batch normalization layer, ReLU activation function layer, 1×1 convolution, 3×3 group convolution, 3×3 depthwise convolution, and Dropout layer. The purpose of the first experiment is to form a baseline for comparison with models after adding various module structures to verify whether each structure can effectively improve model classification performance. In the second through fourth ablation experiments, the standard batch normalization layer, ReLU layer, and Dropout layer are removed from bottleneck modules, and residual connection, forget gate, and update gate structures are sequentially added to bottleneck blocks, gradually forming the SEG block designed in this paper. This design aims to observe the impact of each structure on model performance.

Observing the corresponding accuracy values in Table 5, the accuracy increases significantly from the first to second group and from the third to fourth group, indicating that residual connection and update gate structures help improve model classification accuracy. Comparing experimental data between the second and third groups shows that adding the forget gate does not significantly affect model classification accuracy or parameter count, while the forget gate's attenuation operation effectively reduces model computational load. Comparing experimental data between the third and fourth groups shows that the update gate can significantly improve model classification accuracy while occupying only a small number of parameters and computations. In summary, both the proposed hybrid connection pattern and SEG block structure can effectively enhance model performance.

3.7 Comparison with Methods from Other Literature

To demonstrate the

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.