Research Article

Study on the Grading Model of Hepatic Steatosis Based on Improved DenseNet

Ruwen Yang,1 Yaru Zhou,1 Weiwei Liu,2 and Hongtao Shang2

1First Clinical Medical College, Nanjing University of Chinese Medicine, Nanjing 210004, China
2Affiliated Hospital, Nanjing University of Chinese Medicine, Nanjing 210004, China

Correspondence should be addressed to Hongtao Shang; shanghongtao@njucm.edu.cn

Received 6 January 2022; Revised 8 February 2022; Accepted 23 February 2022; Published 17 March 2022

1. Introduction

The liver, as an important organ of the body, is the key to regulating lipid metabolism in the body. When the body’s lipid metabolism is abnormal, it will lead to lipid accumulation in the liver and liver steatosis. Severe liver steatosis is irreversible and threatens human life and health, while milder liver steatosis can be completely cured by treatment. Therefore, early screening for hepatic steatosis is of clinical importance. At present, the degree of hepatic steatosis is mainly determined by manual grading, in which the imaging physician analyzes the patient’s imaging data to score the grade of hepatic steatosis. This approach suffers from subjective grading bias and low efficiency. In recent years, with the widespread use of deep learning in medicine, new opportunities for intelligent grading of hepatic steatosis have been presented. For example, Qiblawey Yazan and Montalbo Francis Jesmar P. proposed a cascade system to detect, localize, and quantify COVID-19 infection from CT images using encoder-decoder convolutional neural networks (ED-CNNs), U-Net, and feature pyramid network (FPN) [1, 2]. Ben Jabra Marwa used 16 deep learning classifiers to diagnose the validity of COVID-19 from chest X-ray images and found that the combination of deep learning models and integrated classification techniques resulted in the highest confidence level for class 3 classification [3]. Liu Zhenguo et al. identified patients with myasthenia gravis effectively based on the 3D DenseNet deep learning (DL) model of preoperative CT of patients as a complement to the conventional diagnostic criteria for identifying thymoma-associated MG [4]. Riasatian Abtin et al. used a DenseNet topology with four dense blocks, fine tuned and trained with different structures to propose a KimiaNet histopathology image recognition model, and tested KimiaNet using three publicly available datasets of TCGA, endometrial cancer images, and colorectal cancer images to verify the effectiveness of the model [5]. The model has certain search and classification performance when used for image representation. Considering the correlation of multilead ECG, systematically mining the correlation of interlead signals, and enhancing the multiplexing of feature information between interlead and intralead signals by using the dense connection of DenseNet, Xiong Peng et al. proposed a novel multilead myocardial infarction localization method based on a densely connected convolutional network (DenseNet), which automatically captures valid myocardial infarction,
improves its recognition rate, and can be introduced into clinical practice to assist in the diagnosis of myocardial infarction [6]. Albahli Saleh et al. proposed three different BiT models: the DenseNet, InceptionV3, and Inception-ResNetV4 for the diagnosis of coronavirus pneumonia patients by X-ray chest radiographs, and the results showed that the pretrained DenseNet model had the highest classification efficiency of 92%, the accuracy of Inception V3 was 83.47% and that of Inception-ResNetV4 was 85.57%. This is sufficient to show the advantages of the DenseNet algorithm [7]. Wang Gaihua et al. proposed an image classification model with a residual attention mechanism based on the improved DenseNet, extracting image features from the training set, which can improve the accuracy of DenseNet algorithm by 8.89% [8]. It can be seen that deep learning has been effective in identifying various diseases in medicine, and it has certain auxiliary functions for physicians to diagnose diseases. Therefore, based on the above research studies, this paper combines the DenseNet algorithm of deep learning with transfer learning to propose a migration learning-based method for grading hepatic steatosis with the improved DenseNet model.

2. Basic Methods

2.1. Introduction to DenseNet Algorithm. The DenseNet algorithm is a neural network that uses feature combinations and set bypasses to improve the performance of the network, which enables the summation of the features of the two pathways before and after the block by means of dense connections [9], which in turn improves the reuse of the features by the network. One L-layer DenseNet network includes \( L(L + 1)/2 \) connections to ensure that model short-circuit values occur in blocks. However, to ensure no short circuiting between blocks, the model adds a pooling layer to approximately reduce the parameters, while ensuring a relatively small model size. The DenseNet model block is shown in Figure 1. The input of layer \( i \) is \( X_i \), which can be expressed as

\[
X_i = H_i([X_0, X_1, \ldots, X_{i-1}]),
\]

where \( H \) is the nonlinear transform, usually a combination of BN + ReLU + Conv\((3 \times 3)\); \( X_0 \) and \( X_{i-1} \) are all layers before the \( i \)-th layer; \([ \ ]\) is the stitching, indicating that all outputs from \( X_0 \) to \( X_{i-1} \) are stitched.

2.1.1. DenseNet Algorithm Improvements. The DenseNet model extracts features by training from scratch, which tends to lead the model fall into local optimum. In addition, to obtain the best prediction, the model often needs a large amount of data support, which leads to a deepening network, and a too deep network will reduce the learning ability of the model, which in turn is prone to overfitting. At the same time, the radio image used for liver steatosis grading is usually small compared to conventional image datasets so that the model is less effective in classification recognition. To solve the above problems, the DenseNet model is improved in this paper. According to the literature [10, 11], if two domains have commonality, their "knowledge" can be transferred; i.e., the knowledge learned from one domain can be used in the other domain. Therefore, this paper improves the DenseNet model by combining transfer learning. In this paper, we improve the training accuracy of the model by transferring the pretrained model from the world’s largest image recognition library (ImageNet dataset) into the DenseNet model [12].

When training a DenseNet model based on transfer learning, the amount of data (batch-size) fed into each iteration is only a small fraction of all the data, and the larger the batch-size, the less time it takes to train the model for one round (epoch). After increasing the batch size \( n \) times, the time per batch is \( \sqrt{n} \). Considering the existence of a local minimum in the design training, the objective of the training is to minimize the softmax function. During the training process, the learning rate determines the update step of the iterations, and if its value is too large, it tends to cause the loss function not to converge, too small to fall more slowly. The best value can be selected by observation.

The regular model gradient descent update weight is as follows [13]:

\[
w' = w - a \cdot \nabla c,
\]

where \( w' \) and \( w \) denote the connection weights after and before the update, \( a \) denotes the learning rate, and \( \nabla c \) denotes the backpropagation gradient. After adding the momentum, the above equation can be rewritten as

\[
v' = \text{momentum} \cdot v - a \cdot \nabla c,
\]

\[
w' = w + v',
\]

where \( v \) denotes the iterative gradient cumulative information and momentum is the momentum coefficient, which usually takes the value of 0.9 to 0.99 [14].

3. Deep Learning for Liver Steatosis Grading Practice

Based on the above analysis of the DenseNet model related to transfer learning, the specific method of using deep learning for liver steatosis grading study in this paper is as follows:

(1) Dataset construction. All patients’ clinical information is concealed, the grade of hepatic steatosis is labeled and used as a tag, and the dataset with their corresponding ROI is constituted.

(2) Data preprocessing. To improve the convergence speed of the network, the datasets are dealt with normalization, the patient MRI image pixels are normalized to [0,1] without changing the stored information of images, and the training set and test set are divided in a certain ratio [15].

(3) Model construction and training. Based on the basic structure of DenseNet, a DenseNet model is built and transfer learning feedforward and backpropagation algorithms are trained on low-level weights and
high-level weights, respectively, to identify the features of the discovered structures and specific images in the images. In this paper, 1.28 million natural images from the ImageNet dataset are selected for pretraining the model.

(4) Model fine tuning. Model fine tuning usually includes by training the whole network, freezing the convolutional base, and freezing some layers to train some layers [16]. Considering the difference in data volume between the experimental dataset and the ImageNet dataset on the hepatic steatosis study in this paper, it was decided to choose the approach of freezing some layers to train some layers for model fine tuning[17]. After the model is pretrained, the fully connected layer is removed, 2 dense blocks are frozen, and 1 untrained dense block and the fully connected layer are added immediately afterwards, with the aim of extracting and classifying advanced features, which then connected by batch normalization and pooling layers [18]. Ultimately, the model structure used in this paper for liver steatosis grading study is shown in Figure 2. In the figure, Dense Block1, Dense Block2, and Dense Block3 are densely connected in 4, 4, and 32 layers, respectively.

4. Simulation Experiments

4.1. Experimental Environment Setup. The experiments were conducted on an Nvidia GeForce RTX 2080 Ti GPU, implemented through the Keras deep learning framework, and programmed in Python 3.6.

4.2. Data Source and Preprocessing. The data of this experiment were obtained from abdominal MR imaging data and clinical information of 50 patients from June to July 2020 in a hospital in Beijing. Among them, hepatic fat grade grading was done independently by two professional imaging physicians based on patient imaging data, and hepaticocellular steatosis was divided into four degree scores, noting hepaticocellular adipocyte deformation 0–5% as 0, 5–33% as 1, 34%–66% as 2, and more than 66% as 3, respectively [19]. The statistics of patients’ hepatic steatosis grade shows that there were 38 patients with grade 0, 24 male patients and 14 female patients; 12 patients with grade 1, 6 male patients and 1 female patient; there were no patients with grade 2 or 3.

Figure 2 shows an example of an MR abdominal mDixon imaging slice of a patient. Six square ROI regions were set for each MR sequence of DICOM images under the premise of avoiding large blood vessels, focal liver lesions, and significant liver artifacts [20]. Four of the ROIs were located in the parenchyma of the right lobe of the patient’s liver, i.e., segments V, VI, VII, and VIII of the liver, and the other two ROIs were located in the parenchyma of the left lobe of the patient’s liver, i.e., segments II and III of the liver. Each region of the liver parenchyma has 16 * 16 pixels.

Considering the small amount of experimental data, only 300 ROIs of 50 patients, it is difficult to meet the demand of deep learning data volume, so the experiment uses panning, rotation, mirroring, and other enhancement processing on the data [21]. In addition, since the difference in the number of data between samples with liver fat deformation as grade 0 (38 cases) and grade 1 (12 cases) is more obvious, if directly input into the depth model, it will easily lead to the neglect of the minority class samples, which are of important research value in medical image analysis. Therefore, to avoid the effect of unbalanced data samples on the results, the experiments combined the actual number of minority class samples and majority class samples and used the method of oversampling minority class samples to deal with unbalanced data samples [22, 23]. The 648 ROIs with grade 1 hepatic steatosis were amplified twice with 2-fold data to obtain a total of 2,592 cases of grade 1 hepatic steatosis ROI data. Finally, 2,052 cases of grade 0 ROI data and 2,592 cases of grade 1 ROI data of hepatic steatosis were obtained in this experiment, where 3,766 and 878 ROIs were used for model training and testing, respectively, and Adam was selected as the optimization algorithm during the modeling process.

4.3. Evaluation Indexes. In this experiment, precision, sensitivity, specificity, and AUC were chosen as the indexes to evaluate the model performance. Precision reflects the probability that the model predicts correctly for all samples and is calculated as in equation (5). Sensitivity is a measure of the probability that the model predicts correctly for positive class samples and is calculated as in equation (6).
The specificity is a measure of the probability that the model correctly predicts the negative class samples and is calculated as in equation (7). The higher the values of precision, sensitivity, and specificity, the better the model performance. AUC, the probability of predicting a positive case before a negative case, is an important index of the predictive effectiveness of a dichotomous classification model. The closer its value is to 1, the better the model performance is indicated [24].

\[
\text{precision} = \frac{TP}{(TP + FP)} \quad (5)
\]

\[
\text{sensitivity} = \frac{TP}{P} \quad (6)
\]

\[
\text{specificity} = \frac{TN}{N} \quad (7)
\]

where \(TP\) denotes true positive; \(FP\) denotes false positive; \(P\) denotes all positive; \(TN\) denotes true negative; \(N\) denotes all negative. The accuracy and the loss function were selected to evaluate the metrics for evaluating the change of the model during training and testing. The accuracy rate reflects the probability of correct prediction during the model iteration and is calculated as in equation (8). The closer its value is to 1, the higher the model prediction accuracy is. The loss function reflects the difference between the predicted label and the true label, and is calculated as in equation (9). The closer its value is to 0, the better the model prediction is, and the closer the predicted value is to the true value [25].

\[
\text{accuracy} = \frac{TP + TN}{P + N} \quad (8)
\]

\[
\text{loss} = [y \log \bar{y} + (1 + y) \log (1 - \bar{y})] \quad (9)
\]

In equation (9), \(y\) and \(\bar{y}\) denote the true label and the predicted label, respectively.

4.4. Experimental Results

4.4.1. Model Validation. To verify the effectiveness of the proposed model transfer learning for model improvement, the experiments compare the training error of the model before and after transfer learning with the test set accuracy, and the results are shown in Figure 3, where \(a\) and \(b\) are the model training error and test set accuracy before transfer learning, with the number of iterations, respectively, and \(c\) and \(d\) are the model training error and test set accuracy after transfer learning, with the number of iterations, respectively. As can be seen from the figure, before transfer learning, the model reached convergence at about 100,000 iterations and its accuracy on the test set was 80%; after transfer learning, the model converged at about 30,000 iterations and its accuracy on the test set was 83.4%. It shows that transfer learning can improve the model convergence speed and prediction accuracy.

To further validate the effectiveness of transfer learning, the confusion matrix of the model’s prediction of each grade of hepatic steatosis after transfer learning was experimentally analyzed, as shown in Figure 4, where the horizontal and vertical coordinates are the sample prediction and the true grade, respectively, and the diagonal line is the proportion of correct predictions. As can be seen from the figure, the proposed model has a high accuracy in predicting the grade of hepatic steatosis, which is more than 80%.

In addition, the experiments also compared the prediction effects of the proposed model with other models before and after transfer learning, and the results are shown in Table 1. As can be seen from the table, the accuracy of the model after transfer learning is improved in different degrees on the test set compared with the model before transfer learning; the proposed model has a higher accuracy of 83% on the test set compared with other transfer learning models, and the smaller the model size is, the shorter the time to predict a single image, which shows that the proposed model has certain superiority.

4.4.2. Model Prediction Results. To verify the validity of the proposed model for liver steatosis grading practice, the accuracy, sensitivity, specificity, and AUC values of the model on the training and test sets were collected experimentally, and the results are shown in Table 2. As shown in the table, the accuracy of the model on the training and test sets was more than 85%, the sensitivity was more than 94%, the specificity was about 80%, and the AUC value was more than 0.8, which is closer to 1. This indicates that the proposed model has good predictive performance and can predict the grade of hepatic steatosis of patients more accurately. The sensitivity and specificity indicate that the model has a strong ability to detect positive cases, indicating that the model can detect hepatic steatosis grade 1 more reliably and has some practical application value.

Figure 5 shows the changes of loss and accuracy curves on the training and testing sets during the iterations of the model. From the figure, it can be seen that with the increase of iterations, the loss and accuracy curves gradually leveled off and reached a stable state after 400 iterations; the model did not show any overfitting phenomenon during the whole training and testing process, which indicates that the proposed model exhibits good prediction performance through transfer learning.

4.4.3. Statistical Analysis. Spielman correlation analysis was performed using MATLAB R2018b on the obtained clinical information, such as gender and age of the patients, and the grade of hepatic steatosis, and the results are shown in Table 3, where \(P < 0.05\) indicates a statistically significant difference, and * indicates a significant correlation at the 0.05 level. The table shows that the \(p\) value of pancreatic steatosis grade and hepatic steatosis grade was 0.003, and the correlation between them was significant; the \(p\) value of other clinical information such as patient age and gender and hepatic steatosis grade was larger than 0.05, indicating that the correlation between them was not significant.

Correlation analysis between clinical information and hepatic steatosis grade in patients of different genders
revealed significant correlations between hepatic steatosis grade and metabolic syndrome and pancreatic steatosis grade in female patients $P = 0.007$ and $P = 0.002$, while there was no correlation between hepatic steatosis grade and its clinical information in male patients. The change in liver fat content with age of the patients was analyzed, and the results are shown in Figure 6. As can be seen from the figure, there was no correlation between liver fat content and the age of
Table 1: Comparison of prediction results of different models.

<table>
<thead>
<tr>
<th>Network name</th>
<th>Testing set accuracy</th>
<th>Nontransferable learning (%)</th>
<th>Transfer learning (%)</th>
<th>Model size (MB)</th>
<th>The time of predicting single image(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResNet52</td>
<td>68.34</td>
<td>70.33</td>
<td>235</td>
<td>0.2447</td>
<td></td>
</tr>
<tr>
<td>VGG16</td>
<td>77.25</td>
<td>79.48</td>
<td>521</td>
<td>0.0630</td>
<td></td>
</tr>
<tr>
<td>VGG19</td>
<td>77.31</td>
<td>80.27</td>
<td>561</td>
<td>0.0704</td>
<td></td>
</tr>
<tr>
<td>SqueezeNet</td>
<td>73.09</td>
<td>77.36</td>
<td>2.77</td>
<td>0.0072</td>
<td></td>
</tr>
<tr>
<td>GoogleNet-inception V1</td>
<td>78.42</td>
<td>80.46</td>
<td>39.4</td>
<td>0.0221</td>
<td></td>
</tr>
<tr>
<td>DenseNet 16</td>
<td>80.27</td>
<td>83.46</td>
<td>101</td>
<td>0.2721</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Model performance.

<table>
<thead>
<tr>
<th>Items</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training sets</td>
<td>88.49</td>
<td>95.44</td>
<td>81.6</td>
</tr>
<tr>
<td>Test sets</td>
<td>85.79</td>
<td>94.55</td>
<td>79.82</td>
</tr>
</tbody>
</table>

Figure 5: Model performance iteration curve.

Table 3: Analysis results of clinical information and hepatic steatosis grade of patients.

<table>
<thead>
<tr>
<th>Clinical information</th>
<th>Correlation coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.003</td>
<td>0.959</td>
</tr>
<tr>
<td>Gender</td>
<td>0.035</td>
<td>0.570</td>
</tr>
<tr>
<td>Pancreatic steatosis grade</td>
<td>0.702 *</td>
<td>0.003</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td>0.890</td>
<td>0.147</td>
</tr>
</tbody>
</table>
the patients, but there was a peak between the ages of 50 and 60 years for women and one peak each between the ages of 40 and 50 years and 70 to 80 years for men.

5. Conclusion

In summary, applying deep learning to liver steatosis time can achieve automated and more accurate grading, and by introducing transfer learning into the deep learning DenseNet model, the number of model iterations can be significantly reduced and the model convergence speed and prediction accuracy can be improved, which has some potential application value. Applying it to the practice of liver steatosis grading, the accuracy of the proposed model on the training and test sets reached more than 85%, the sensitivity exceeded 94%, the specificity was about 80%, and the AUC value reached more than 0.8, which is closer to 1. It has good predictive performance and can detect liver steatosis grade 1 more accurately and reliably, which has some practical application value. Statistical correlation analysis shows that hepatic steatosis correlates significantly with pancreatic steatosis and has some correlation with metabolic syndrome; hepatic steatosis was correlated with age and metabolic syndrome in women only. However, due to the limitations, there are still some shortcomings in this paper; the amount of patient data is too small, the prediction accuracy can be improved, which has some practical application value. Applying it to the practice of liver steatosis grading, the accuracy of the proposed model on the training and test sets reached more than 85%, the sensitivity exceeded 94%, the specificity was about 80%, and the AUC value reached more than 0.8, which is closer to 1. It has good predictive performance and can detect liver steatosis grade 1 more accurately and reliably, which has some practical application value. Statistical correlation analysis shows that hepatic steatosis correlates significantly with pancreatic steatosis and has some correlation with metabolic syndrome; hepatic steatosis was correlated with age and metabolic syndrome in women only. However, due to the limitations, there are still some shortcomings in this paper; the amount of patient data is too small, the preliminary results obtained need to be further validated, and among the 50 patients with hepatic steatosis, there are no patients with grade 1 or higher, which leads to the accuracy and generalization ability of the model to be demonstrated. Therefore, the amount of hepatic steatosis imaging data will be further collected and expanded in the next study.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References


