Utility of convolutional neural network-based algorithm in medical images for liver fibrosis assessment

Wei Wei¹, Xu Yang²

¹Clinical Epidemiology and EBM Unit, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China; National Clinical Research Center for Digestive Diseases, Beijing 100050, China;
²School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China.

To the Editor: Liver fibrosis is the critical stage leading to hepatic dysfunction and might be important in the progress to portal hypertension, biliary cirrhosis, and hepatocellular carcinoma. Therefore, the accurate assessment of liver fibrosis remains a clinical concern for physicians. In recent researches, liver biopsy has been considered the golden standard for the diagnostic and assessment of liver fibrosis for many years. However, due to its invasive practice, sampling variability, and inter- and intra-assessor variability of pathological interpretations, liver biopsy is not well received in the patients and clinical physicians in many circumstances. As a result, multiple non-invasive detections of hepatic fibrosis assessment have been developed as an alternative strategy for liver biopsy as they could offer a convenient operation and acceptable diagnostic accuracy through medical imaging information. Early detection of hepatic fibrosis through clinical imaging could reduce liver failure and prevent its progression. The interpretation and analysis of medical imaging information are usually performed by clinical experts. Benefit from the rapid development of computer-aided diagnosis, especially the improved algorithm of deep learning in artificial intelligence, physicians could extract more accurate assessment for clinical decisions for diagnosis and treatment.¹¹

The deep learning convolutional neural network (CNN) has been particularly preferred for medical imaging processing in hepatic fibrosis assessment, including medical image segmentation, clinical classification, and prediction. CNN has also been demonstrated to be robust against data heterogeneity and encouraged for the assessment of different liver fibrosis stages.²² CNN has shown high diagnostic accuracy in image classification and received popularity in many other medical fields.

In our research, we aim to investigate the basic techniques of CNN and the recent CNN algorithm for hepatic fibrosis in medical imaging to improve diagnostic accuracy, which includes imaging feature extraction, convolved operation, normalization, and rectified linear units (ReLU) operation. In addition, we review the current studies using the CNN algorithm for evaluating the hepatic fibrosis stage and discuss the potential deep learning algorithm in medical imaging for hepatic fibrosis assessment.

In our study, we searched the following databases for related studies: MEDLINE, EMBASE, Chinese Biomedical Literature Database, WANFANG, and CNKI, covering the publish period between January 1, 1966, and January 1, 2020. The retrieval strategy was as follows: “(liver fibrosis OR hepatic fibrosis) AND (convolutional neural network OR CNN OR DCNN) AND (accuracy OR sensitivity OR specificity OR ROC OR AUROC) AND (prediction OR diagnostic OR diagnosis). The abstracts and full text of researched studies were assessed by Wei and Yang. A third investigator would be invited to resolve conflicts. The inclusion criteria were: (1) Patients were aged from 18 years to 65 years at baseline. (2) CNN algorithm was used for the assessment of the liver fibrosis stage. (3) The accuracy of the diagnostic model was assessed as the following indicators: area under the receiver operator characteristic curve (AUROC), sensitivity (SEN), specificity (SPE), false-positive rate (FPR), or false-negative rate (FNR). The exclusion criteria were: (1) Patients complicated with severe cardiovascular and cerebrovascular disease. (2) Patients suffered from psychological disorders. (3) Patients suffered from malignant digestive tumors.

We have researched eight studies that utilized a CNN-based algorithm for the assessment of various liver fibrosis stages [Supplementary Digital Content, Table 1, http://links.lww.com/CM9/A571]. One research developed an automated framework of CNN classifier for the assessment of liver fibrosis induced by non-alcoholic fatty liver disease (NAFLD) in patients and other seven researches...
were aimed to evaluate the liver fibrosis induced by hepatitis B virus (HBV).

Liu et al.\textsuperscript{3} proposed a computer-aided cirrhosis diagnosis system to diagnose cirrhosis based on ultrasound images, and they proposed a method to extract image features on an ultrasound image using a deep convolutional neural network (DCNN) model. The results showed that the proposed CNN model outperforms the two variants with the highest accuracy of 0.968. Brattain et al.\textsuperscript{4} developed and evaluated an automated framework to check shear wave elastography (SWE) image quality and select a region of interest (ROI). To determine the fibrosis stage, several classifiers including random forests (RF), support vector machine (SVM), and CNN were constructed. The best approach utilized a CNN and corresponding area under the receiver operator characteristic curve (AUROC) was 0.890, compared with the conventional stiffness only based AUROC of 0.740. Byra et al.\textsuperscript{5} constructed the inception-ResNet-v2 DCNN pre-trained on the ImageNet dataset for evaluating liver steatosis in ultrasound images with the highest sensitivity. In the results, the AUROC obtained using the proposed approach was equal to 0.977, higher than the one obtained with the hepatorenal index method of 0.959, and much higher than the gray-level co-occurrence matrix algorithm of 0.893. Wang et al.\textsuperscript{6} evaluated the performance of the newly developed deep learning radiomics of elastography (DLRE) for assessing liver fibrosis stages. The authors enrolled 398 patients with 1990 images and analyzed the receiver operator characteristic (ROC) curves to calculate the optimal AUROC for cirrhosis (F4), advanced fibrosis (F3), and significance fibrosis (F2). The AUROC of DLRE was 0.970 for F4 (95% Confidence interval [CI]: 0.94–0.99), 0.980 for F3 (95% CI: 0.96–1.00), and 0.850 for F2 (95% CI: 0.81–0.89), which were significantly better than other methods except for two-dimensional shear wave elastography (2D-SWE) in ≥F2. Yu et al.\textsuperscript{7} validated a deep learning-based algorithm utilizing pre-trained AlexNet-CNN that would automatically calculate liver fibrosis stages, and AUROC was superior to conventional CNN, non-liver multinomial logistic regression (MLR), SVM, and RF. In AlexNet-CNN, seven hidden layers were constructed to process specific values for momentum and weight decay. Meanwhile, the input and output layers of the traditional AlexNet-CNN network were adapted for liver fibrosis assessment according to the total collagen content, where the processed second harmonic generation hepatic biopsy images were segmented into proper pixel to fit into the network as input images. Based on the above algorithm, several articles focused on quantification assessment of hepatic steatosis in histologic images that aimed to improve diagnostic accuracy compared with invasive biopsy evaluation. Yasaka et al.\textsuperscript{8} investigated DCNN based on dynamic contrast-enhanced computed tomography (CT) images for the evaluation of hepatic fibrosis stage, and the diagnostic accuracy would be evaluated from histopathological information regarding the fibrosis stage. In the DCNN algorithm, the liver CT images fed to the input layer were down-sampled with a max-pooling process and transformed to single-pixel images. The hepatic fibrosis score obtained from DCNN on CT images showed a significant correlation with the fibrosis stage for diagnosing significant fibrosis, advanced fibrosis, and cirrhosis. The hepatic fibrosis could be staged using a deep learning algorithm based on dynamic contrast-enhanced portal phase CT images. Treacher et al.\textsuperscript{9} designed a randomized search of 100 CNN architectures, conducted for parameter optimization to test whether the texture pattern in grayscale elastography images predictive of the shear wave velocity. The results demonstrated the accuracy of experts classifying high versus low fibrosis was 58.3% (standard deviation was 7.6%). Gatos et al.\textsuperscript{10} detected and isolated areas of low and high stiffness temporal stability in SWE images validated by DCNN for different liver fibrosis stage patients. The CNN classification showed improved accuracy (ranging from 82.5% to 95.3%) in SWE images compared with the unmasked ones (ranging from 79.5% to 93.2%) for various chronic liver disease (CLD) stage combinations. The merged SEN, SPE and ROC have been shown in Supplementary Digital Content, Figure 1, http://links.lww.com/CM9/A571.

The convolution layer is the basic foundation of the CNN construction that implements image feature extraction which typically including the combination of linear and non-linear functions. The function of the convolution layer is feature representation by increasing the semantic level features with the depth of layers. In the convolution layer, a kernel consists of a small array of pixel numbers would be applied across the input, which could be learned with backpropagation algorithms.

In the convolution layer, multiple image maps are connected with a plurality of neurons in local regions, which are calculated from the upper-layer image features through the convolution kernel and also the weight matrix. Each local weight matrix would be activated to a non-linear function to solve output value and propagate to subsequent convolutional layers.

In a fully connected layer, the output feature information of the final convolution layer is typically transformed into a flattened one-dimensional array of output numbers calculated through previous convolution operation and pooling operation, then connected to one or more fully connected layers. Once the imaging information abstracted by the kernel and convolution operation is created, they could be mapped into the subset of a fully connected layer toward the final output, which indicates the probable classification for each input imaging feature. The operation of the CNN algorithm has been integrated in Figure 1.

The activation function utilized in the last fully connected layer directly bridges the clinical output, which immensely influences the diagnostic accuracy for the assessment of liver fibrosis. According to the final classification of fibrosis, an appropriate activation function would be applied for normalizing the target fibrosis probabilities ranging from 0 to 1.

The ReLU function could make the output value more than zero even when the input value less than zero. In the
sigmoid function, input variables were fitted by linear regression before the activation function applied, where “a” and “b” represented the intercept and the coefficient, respectively, and the variables transformed to the “s” curve and its value between 0 and 1. In the hyperbolic tangent (tanh) function, the value ranged from $-\text{C}0$ to 1.

Convolutional neural network has outstanding performance in computer-aided diagnoses of liver fibrosis patients, which could automatically extract imaging features and calculate the weights between each neuron through their contribution, making the construction of the network more cost-effective and less tedious. One of the potential advantages of the CNN model is the capability of extracting image features by convolutional process and reducing data dimension by pooling process, especially for deciphering features and pattern recognition based on its architecture structure inspired by visual cortex. One of the challenges for clinicians is the interpretation of vast neurons and nonlinear activation function in CNN algorithm. To solve the above problem, researchers have proposed several techniques to give insight into the image features, which could help clinicians comprehending the CNN model.

**Funding**

This research was supported by the grants from the National Natural Science Foundation of China (No. 91846303) and the Beijing Municipal Administration of Hospitals Incubating Program (No. PX2018071).

**Conflicts of interest**

None.

**References**

1. Japkowicz N, Shah M. Evaluating learning algorithms: A classification perspective. New York: Cambridge University Press; 2011.
2. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018;9:611–629. doi: 10.1007/s13244-018-0639-9.
3. Liu X, Song JL, Wang SH, Zhao JW, Chen YQ. Learning to diagnose cirrhosis with liver capsule guided ultrasound image classification. Sensors (Basel) 2017;17:149. doi: 10.3390/s17010149.
4. Brattain LJ, Telfer BA, Dhyan M, Grajo JR, Samar AE. Objective liver fibrosis estimation from shear wave elastography. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:1–5. doi: 10.1109/EMBC.2018.8513011.
5. Byra M, Styczynski G, Samigielisz C, Kalinowski P, Michalowski Ł, Palaszewicz R, et al. Transfer learning with deep convolutional neural network for liver steatosis assessment in ultrasound images. Int J Comput Assist Radiol Surg 2018;13:1895–1903. doi: 10.1007/s11548-018-1843-2.
6. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. Gut 2019;68:729–741. doi: 10.1136/gutjnl-2018-316204.
7. Yu Y, Wang J, Ng CW, Mo S, Fong ELS, et al. Deep learning enables automated scoring of liver fibrosis stages. Sci Rep 2018;8:16016. doi: 10.1038/s41598-018-34300-2.
8. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiyu S. Deep learning for staging liver fibrosis on CT: A pilot study. Eur Radiol 2018;28:4578–4585. doi: 10.1007/s00330-018-5499-7.
9. Treacher A, Beauchamp D, Quadri B, Fetzer D, Vij A, Yokoo T, et al. Temporal stability assessment in shear wave elasticity images validated by deep learning neural network for chronic liver disease fibrosis stage assessment. Med Phys 2019;46:2298–2309. doi: 10.1002/mp.13521.