

ANALYSIS AND CLASSIFICATION OF HEPATITIS INFECTIONS USING RAMAN SPECTROSCOPY AND MULTI-SCALE CONVOLUTIONAL NEURAL NETWORKS**

Y. Zhao¹, Sh. Tian^{1*}, L. Yu², Zh. Zhang³, W. Zhang¹

¹ Key Laboratory of Software Engineering Technology, Xinjiang University, Urumqi 830000, China; e-mail: aleary@vip.qq.com

² College of Software Engineering at Xin Jiang University, Urumqi 830000, China

³ The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830000, China

Hepatitis infections represent a major health concern worldwide. Numerous computer-aided approaches have been devised for the early detection of hepatitis. In this study, we propose a method for the analysis and classification of cases of hepatitis-B virus (HBV), hepatitis-C virus (HCV), and healthy subjects using Raman spectroscopy and a multi-scale convolutional neural network (MSCNN). In particular, serum samples of HBV-infected patients (435 cases), HCV-infected patients (374 cases), and healthy persons (499 cases) are analyzed via Raman spectroscopy. The differences between Raman peaks in the measured serum spectra indicate specific biomolecular differences among the three classes. The dimensionality of the spectral data is reduced through principal component analysis. Subsequently, features are extracted, and then feature normalization is applied. Next, the extracted features are used to train different classifiers, namely MSCNN, a single-scale convolutional neural network, and other traditional classifiers. Among these classifiers, the MSCNN model achieved the best outcomes with a precision of 98.89%, sensitivity of 97.44%, specificity of 94.54%, and accuracy of 94.92%. Overall, the results demonstrate that Raman spectral analysis and MSCNN can be effectively utilized for rapid screening of hepatitis B and C cases.

Keywords: multi-scale convolutional neural network, hepatitis-B, hepatitis-C, Raman spectroscopy.

АНАЛИЗ И КЛАССИФИКАЦИЯ ИНФЕКЦИЙ ГЕПАТИТА С ИСПОЛЬЗОВАНИЕМ СПЕКТРОСКОПИИ КОМБИНАЦИОННОГО РАССЕЯНИЯ И МНОГОМАСШТАБНЫХ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Y. Zhao¹, Sh. Tian^{1*}, L. Yu², Zh. Zhang³, W. Zhang¹

УДК 535.375.5:616.36-002

¹ Синьцзянский университет, Урумчи 830000, Китай; e-mail: aleary@vip.qq.com

² Колледж программной инженерии университета Синьцзян, Урумчи, 830000, Китай

³ Первая больница Синьцзянского медицинского университета, Урумчи, 830000, Китай

(Поступила 22 января 2020)

Предложен метод анализа и классификации случаев вируса гепатита В (HBV) и С (HCV) с использованием спектроскопии комбинационного рассеяния и многомасштабной сверточной нейронной сети (ММСНС). Проанализированы образцы сыворотки пациентов, инфицированных HBV (435 случаев) и HCV (374 случая), и здоровых людей (499 случаев). Различия между пиками в измеренных КР-спектрах сыворотки указывают на конкретные биомолекулярные различия между группами. Размерность спектральных данных уменьшена за счет анализа главных компонент, к полученным признакам применена нормализация. Полученные признаки использованы для обучения различных классификаторов, а именно ММСНС, одномасштабной сверточной нейронной сети и других традиционных классификаторов. Модель ММСНС имеет лучшие результаты с четкостью 98.89%, чув-

** Full text is published in JAS V. 88, No. 2 (<http://springer.com/journal/10812>) and in electronic version of ZhPS V. 88, No. 2 (http://www.elibrary.ru/title_about.asp?id=7318; sales@elibrary.ru).

ствительностью 97.44%, специфичностью 94.54% и точностью 94.92%. Показано, что спектральный анализ комбинационного рассеяния и ММЧНС могут быть эффективно использованы для быстрого скрининга гепатитов В и С.

Ключевые слова: многомасштабная сверточная нейронная сеть, гепатит В, гепатит С, спектроскопия комбинационного рассеяния.

Introduction. Hepatitis is the second most common disease worldwide with half of the cases in developing countries. The mortality rate of hepatitis is quite high, which can be primarily reduced through early diagnosis and treatment [1]. Most of the hepatitis infections are due to either hepatitis-B virus (HBV) or hepatitis-C virus (HCV) [2]. According to the World Health Organization (WHO), there are 350 and 170 million cases of HBV and HCV chronic infection globally, respectively. Further, the annual number of deaths caused by HBV and HCV are nearly 10,000 and 600,000 [3], respectively. Moreover, people with hepatitis are at a high risk of developing liver cancer [1]. While both types of hepatitis are infectious liver diseases, they can be differentiated based on the key aspects. Firstly, HBV is a DNA-type virus, while HCV is a RNA-type virus [4]. Second, HCV is highly susceptible to mutations. This complicates the treatment and vaccine development for HCV [4].

Raman spectroscopy is a modern scattering-based spectroscopic technique for material composition analysis [5]. This technique is frequently used for non-invasive, real-time diagnosis of cancer and various other infections [6]. For example, Raman spectroscopy has been successfully used in the early diagnosis of dengue fever [7], nasopharyngeal carcinoma [8], esophageal cancer [9], and rectal cancer [10]. In fact, the applicability of Raman spectroscopy in medical diagnosis is based on the disease-specific variations in spectral variables including the locations of the characteristic peaks, intensity of the scattered radiation, and spectral linewidth. The abscissa of the Raman spectrum is the Raman shift, which represents the spectral difference between the incident and scattered light rays. The ordinate of this spectrum is the photon count, which represents the intensity of the scattered light [11].

Over the recent years, deep learning has emerged as one of the key high-performance computer methodologies for extracting information from raw data. Numerous deep learning techniques have been proposed to extract data-driven features and implement classification and prediction tasks in large-scale and big-data problems [12]. Deep learning has especially been applied for solving complex chemical and biological problems in the fields of spectroscopy [13–17], proteomics [18–20], metabolomics [21, 22], and genomics [23, 24]. In particular, convolutional neural networks (CNNs) represent one of the most prominent deep learning architectures. However, conventional single-scale convolution kernels in CNNs miss key details during the information encoding process. Inspired by the InceptionNet [25] architecture, multi-scale convolution kernels are used in this study to extract rich, inconspicuous features from the input Raman spectra. This multi-scale scheme promotes the encoding of more useful Raman spectral features.

We proposed to use Raman spectroscopy in combination with a multi-scale convolutional neural network (MSCNN) to differentiate between HBV, HCV, and healthy samples. Firstly, Raman spectral features are obtained for each sample. Then, principal component analysis (PCA) is applied to reduce the dimensionality of spectral features. Subsequently, mean normalization is performed on the spectral features. Finally, the spectral features are used to train and test a MSCNN architecture. An experiment is designed to compare the performance of deep learning and conventional classifiers. The deep learning classifiers include a single-scale CNN, MSCNN, and MSCNN with a long short-term memory (LSTM). The conventional classifiers include the naive Bayes classifier, *k*-nearest neighbor (KNN) classifier, logistic regression, random forests, decision trees, support vector machines (SVMs), and gradient boosting decision trees (GBDTs).

Materials and methods. *Serum sample collection.* For the experimental evaluation of the proposed method, we collected 435, 374, and 499 serum samples of hepatitis-B patients, hepatitis-C patients, and healthy subjects, respectively. The diagnosis and sample collection were done at the First Affiliated Hospital of Xinjiang Medical University. The serum samples of the HCV-infected patients passed the reverse transcriptase polymerase chain reaction (RT-PCR) test. The RT-PCR method was used to quantitatively detect the HCV RNA. For each HBV- and HCV-infected patient, we collected 5 mL of fresh blood after 8 h of fasting. After full coagulation of the blood sample, components such as fibrinogen and blood cells were removed by a high-speed centrifuge at a speed of 3000 rpm. Then, a serum volume of 500 μ L was collected in an Eppendorf tube and stored at -80°C . Before the examination, the specimen under test was placed in an environment of 22°C for 30 min. After sufficient dissolution, 30 μ L of the serum sample was captured in a capillary glass tube for spectral analysis. The age and gender information of the study sample is shown in Table 1.

TABLE 1. Demographic Data of the Study Population

Group	Control ($n = 499$)	HBV ($n = 435$)	HCV ($n = 374$)
Age (years, s.d)	19–86	20–79	22–89
Gender (%F)	44.9	47.4	50.0

Raman spectroscopy. The source laser beam was focused on the surface of a droplet by a 50× microscope objective (numerical aperture = 0.5) with a spectral resolution of 5 cm⁻¹ and a recording time of 1.0 s (LabRAM HR Evolution Raman Spectrometer, HORIBA Scientific Ltd.). The spectral data resolution and Raman shift were measured three times and averaged, and the corresponding average values were 5 and 300–3000 cm⁻¹, respectively.

Preprocessing and analysis of spectral data. Since the raw Raman spectra obtained from the serum samples contained significant fluorescence background and noise, spectral preprocessing was necessary [26–28]. The preprocessing steps were as follows: the fluorescent background was removed; spectral smoothing was applied; PCA was used to reduce the feature dimensionality. For background subtraction, the Vancouver Raman algorithm based on a fifth-order polynomial was used to fit all the autofluorescence backgrounds of serum. The Savitzky–Golay filter was used to smooth the spectra of all the samples, and the fifth-order polynomial was subtracted to correct the baseline [29]. Finally, comparative experiments were conducted to select the appropriate samples.

Evaluation criteria. For the experiments, we divided the collected data into training and test sets with sizes in the ratio of 4:1. For performance evaluation, a confusion matrix was used to numerically quantify the misclassification among the pairs of classes [30]. Specifically, this matrix described the confusion between the real and output classes, where the non-diagonal element in the i -th row and j -th column indicated the number of samples in the i -th class that were wrongly recognized by the classifier as samples of the j -th class. The number of the correctly classified samples of the i -th class was stored as the i -th diagonal element in the confusion matrix [31]. The following performance measures can be derived from the confusion matrix.

Confusion matrix		Predicted class		
		0	1	2
True class	0	a	b	c
	1	d	e	f
	2	g	h	i

$$\text{Precision}_{\text{class0}} = \frac{e + i}{d + e + g + i},$$

$$\text{Specificity}_{\text{class0}} = \frac{a}{a + d + g},$$

$$\text{Sensitivity}_{\text{class0}} = \text{Recall}_{\text{class0}} = \frac{a}{a + b + c}.$$

Accuracy: Ratio of the overall number of the correctly classified samples to the total number of samples.

Precision: Number of relevant positive samples among the samples predicted to be positive.

Sensitivity: Ratio of the truly positive samples that have been correctly classified.

Specificity: Ratio of the truly negative samples that have been correctly classified.

F1-Score: This a measure of the classification accuracy based on a combination of the precision and recall. The value of this measure ranges from 0 (worst performance) to 1 (best performance).

Advanced spectral data processing. PCA-based dimensionality reduction. PCA is a dimensionality reduction method based on linear transformations. Essentially, the input features are projected to several coordinates for maximizing the feature variance by coordinate rotation and obtaining a new data representation that is more aligned with principal variations. Dimensionality reduction is achieved by exploiting multicollinearity [32]. The PCA-based dimensionality reduction algorithm can be outlined as follows:

Input: A Raman spectral sample set $X = \{x_1, x_2, \dots, x_m\}, x_m \in R^d$.

Steps: 1) Subtract the data mean: For the Raman intensity corresponding to each Raman shift, subtract

the mean of the j -th feature from each i -th sample, $x_i^{(j)} = x_i^{(j)} - \mu_j$ where $\mu_j = \frac{1}{m} \sum_{k=1}^m x_k^{(j)}$. The mean values of the features can be combined into a mean vector $\mu = (\mu_1, \mu_2, \dots, \mu_d)$, $\mu \in R^d$.

2) Calculate the covariance matrix of the Raman spectral data $\Sigma = XX^T$.

3) Perform an eigenvalue decomposition of the covariance matrix $\Sigma = XX^T$. Use the eigenvectors $u_1, u_2, \dots$ corresponding to eigenvalues $\lambda_1 > \lambda_2 > \dots$ to form the transformation matrix U .

4) Generate features of reduced dimensionality, $Z = \{Z_1, Z_2, \dots, Z_m\}$, $Z_i = U^T x_i$, $i = 1, 2, \dots, m$.

These features represent an approximation of the original features, $D = \{U_{Z_i} + \mu\}$, $i = 1, 2, \dots, m$.

Output: $Z \in R^{l \times m}$; $D \in R^{d \times m}$.

Mean normalization. Normalization or scaling of features [33] ensures the same range of values for each feature. Normalization can be formulated as follows: $x = \frac{\text{value} - u}{\max}$. This operation leads to numerical

stability and helps in reaching the optimal solution. Indeed, normalization causes the gradient to always move in the direction of the minimum objective value. This leads to the acceleration of convergence and faster attainment of the global minimum [34].

Model construction. *Multi-scale convolution neural network.* MSCNNs were used as the learning framework in this work. An MSCNN consists of three parts: multi-scale feature extraction layers, multi-scale feature fusion layers, and the dimensionality reduction layers. The architecture of MSCNN is shown in Fig. 1.

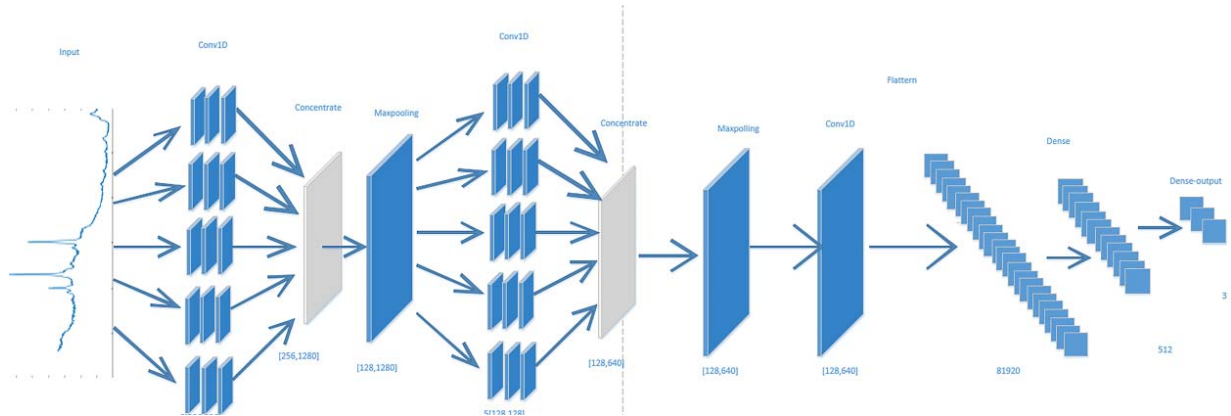


Fig. 1. Architectural model of MSCNN.

Multi-scale feature extraction layers. Multi-scale feature extraction layers extract data characteristics from multiple scales. This layer is mainly composed of parallel convolution kernels of multiple scales. The characteristics of the convolution kernels at different scales are different. In this paper, we apply this feature extraction scheme with five scales or sizes of convolution kernels. The first multi-scale feature extraction layer has five kernel scales of 2×1 , 3×1 , 4×1 , 5×1 , and 10×1 . The second multi-scale feature extraction layer also has five kernel scales of 1×1 , 3×1 , 5×1 , 7×1 , and 9×1 . There are various methods for extracting feature scales. Too few feature scales can lead to underfitting, while too many feature scales can cause overfitting. Dropout and batch normalization modules are added after each convolutional layer. The dropout module can prevent overfitting. Batch normalization significantly speeds up the training of the deep learning model $x f_1(x) = \delta l_i (W_{l_i} * x + B_{l_i}) W_{l_i} B_{l_i} * \delta l_i$ $\delta l_i(x) = \max(0, x)$ [35]. $f^1(x) = \sum_{i=1}^m f_1(x) = \sum_{i=1}^m \delta l_i (W_{l_i} * x + B_{l_i})$.

Let the input of each multi-scale feature extraction layer be x , where the model contains m convolutional layers, and the sizes of convolution kernels between the layers are different. Then, the i -th convolutional layer of the first multi-scale feature extraction module can be modeled as

$$f_i(x) = \delta l_i (W_{l_i} * x + B_{l_i}),$$

where W_{l_i} and B_{l_i} represent the weight and offset of the convolution kernel in the i -th convolutional layer

in the first multi-scale fusion module, respectively; * indicates the convolution operation; δ indicates the i -th convolution after the implementation of rectified linear units (ReLU) δl_i , which are employed in this method and defined as

$$\delta l_i(x) = \max(0, x),$$

where x is the ReLU input.

Multi-scale feature fusion layers. A multi-scale feature fusion layer fuses multiple features after multi-scale feature extraction [35]. The mathematical expression for the output of such layer is

$$f^1(x) = \sum_{i=1}^m f_i(x) = \sum_{i=1}^m \delta l_i(W_{l_i} * x + B_{l_i}),$$

where x represents the input of the multi-scale feature extraction layer, $f^1(x)$ represents the output of the first multi-scale feature fusion layer, and the parameters W_{l_i} and B_{l_i} represent the weight and offset in the dimensionality reduction layer, respectively. The operator δl_i represents the operation of the ReLU nonlinear activation unit.

Dimensionality reduction layer. The dimensionality reduction layer is composed of a convolution layer with a kernel size of $1*1$. The number of convolution kernels in the convolutional layer is less than that of the channels in the multi-feature fusion layer. The $1*1$ kernel size can maintain all the feature information of the multi-feature fusion layer. Due to the single-channel, the multi-scale feature map obtained by the final convolution is also reduced. This leads to a reduced dimensionality. The output of the dimensionality reduction layer of the first multi-scale feature fusion module is expressed as follows:

$$F^1 = \delta(W_{l_{m+1}} * f^1(x) + B_{l_{m+1}}),$$

where x represents the input of the multi-scale feature extraction layer, and $f^1(x)$ represents the output of the first multi-scale feature fusion layer. The parameters $W_{l_{m+1}}$ and $B_{l_{m+1}}$ represent the weight and offset in the dimensionality reduction layer, respectively, and δ represents the ReLU nonlinear activation unit of the dimensionality reduction layer.

The outputs of the multi-scale feature fusion modules of layers other than the first layer ($l > 2$) are expressed as follows:

$$F^l = \delta(W_{l_{m+1}} * \sum_{i=1}^m \delta l_i(W_{l_i} * F^{l-1} + B_{l_i}) + B_{l_{m+1}}),$$

where F^{l-1} represents the output of the previous multi-scale feature fusion module, and the parameters $W_{l_{m+1}}$ and $B_{l_{m+1}}$ represent the weight and offset of the convolution kernels in the dimensionality reduction layer, respectively. The parameters W_{l_i} and B_{l_i} represent the weight and offset of the convolution kernels in the i -th convolutional layer of the first multi-scale feature fusion module, respectively.

LSTM architecture is often used in natural language processing, especially in sentiment analysis. Further, it has been used in classification models that fuse multiple features [36] as well as in analog classification problems. However, the results show that the accuracy of LSTM is 4.05% less than that of MSCNN. LSTM increases the training time, and the relative advantages of MSCNN are obvious.

Results and discussion. Raman spectral analysis. As a type of fingerprint, the Raman spectra of serum can be used to detect changes in biological molecules, such as proteins, nucleic acids, and lipids, of different organisms and interpret the physiological fluctuations of the body to diagnose diseases.

As shown in Fig. 2, after the baseline correction, the values of Raman spectral intensity for HBV-infected patients are generally higher than those for healthy subjects, while the values for HCV-infected patients are essentially lower than those for healthy subjects. The Raman shift is 1154.58 cm^{-1} . Further, the peak intensities of the three spectra are 2910, 2731, and 2213 a.u. for the HBV, healthy, and HCV subjects, respectively. Through baseline correction, the differences between the principal components (PCs) of the three spectral classes can be better explained. The three Raman spectra are quite similar in shape. However, there are intensity differences at Raman shifts of 509, 875, 1018, 1154, 1155, 1444, 1595, 2660, and 2850 cm^{-1} .

TABLE 2. Sample Raman Shifts and their Assignments

Raman shift, cm^{-1}	Assignment
509	S-S disulfide stretching band of collagen v (S-S) gauche-gauche-gauche (amino acid cysteine) [37]
875	Antisymmetric stretching vibration of the choline group $\text{N}^+(\text{CH}_3)_3$, characteristic for phospholipids [37]
1018	Stretching C-O ribose [38]
1154	Deformational mode of carbohydrates [39]
1155	Spectral signature for the CO group of ribose [40, 41]
1595	Benzoquinone [42]
2850	Lipids, fatty acids [42]
860–870	Possible changes in polysaccharide structure are observed in this region [43]
960–980	N(C–C) stretching (probably in amino acids) [43]
1000–1030 cm^{-1}	This region allowed for a complete separation between HBV- and HCV-infected and normal blood serum spectra [43]
1400–1460	This region allowed for a complete separation between HBV- and HCV-infected and normal blood serum spectra [43]
1400–1600	C-120 protein in human blood serum [43]
1500–1600	Amide-I bond [43]

Table 2 lists the Raman shifts and the associated causes for the samples. The Raman peaks at 509 and 2660 cm^{-1} correspond to cysteine and methionine, respectively [44, 45]. Since the liver is crucial for amino acid metabolism, HBV causing liver damage also leads to disorders in the amino acid metabolism, thereby affecting the amino acid profile of serum [46]. The Raman peak at 875 cm^{-1} corresponds to the antisymmetric stretching vibration of the choline group $\text{N}^+(\text{CH}_3)_3$, which has the characteristics of phospholipids, phosphatidylcholine, and sphingomyelin [47]. The peak at 1444 cm^{-1} in the serum spectrum is assigned to CH_2 wagging of cholesterol and lipids, which exhibits a higher intensity in the HBV infection group. Bremer et al. found that the HBV infection depends on the presence of cholesterol in the viral envelope and is essential for the entry process of the HBV [48]. Li et al. demonstrated that the HBV aggravates the accumulation of hepatic cholesterol, which affects the content of serum cholesterol [49]. The main characteristic peaks in the Raman spectrum that correspond to HBV are located at 1018, 1154, 1595, and 2850 cm^{-1} . Among them, the peak at 1018 cm^{-1} is related to a ribose (DNA), while the one at 1154 cm^{-1} corresponds to the deformation of carbohydrates. The peak at 1595 cm^{-1} is relatively strong, which may be attributed to an early liver injury that occurred due to a viral infection, for which benzoquinone is an important biomarker. On the other hand, the Raman peak at 2850 cm^{-1} corresponds to lipids and fatty acids. The differences in the strength of these peaks may be due to the imbalance of liver metabolism; thus, they can be used as a characteristic peak in the diagnosis of HBV [50].

The Raman peak at 1155 cm^{-1} is a spectral signature for the CO group of ribose, which is already reported as the most reliable marker for the HCV detection in serum [51]. Therefore, our results validate the previous reports on this Raman band for detecting HCV infection. The presence of the positive difference bands in the range of $1200\text{--}1400 \text{ cm}^{-1}$ indicates that the base stacking of A-RNA sequences is significantly perturbed under interaction with HCVc-120 protein [52].

Furthermore, the $1400\text{--}1600 \text{ cm}^{-1}$ region in the Raman spectrum of HBV and HCV is due to the presence of C-120 protein in human blood serum, which exhibits a strong broad amide-I bond centered near $1500\text{--}1600 \text{ cm}^{-1}$. Compared to the normal samples, the protein and lipid structures of HBV and HCV specimens are different. Spectral changes are observed in the protein bands, amide I ($1560\text{--}1590 \text{ cm}^{-1}$), amide 3, and N(C–C) stretching (possibly in amino acids, $960\text{--}980 \text{ cm}^{-1}$) [40, 53]. In addition, potential variations in the polysaccharide structure are noticed in the range of $860\text{--}870 \text{ cm}^{-1}$. The analysis of band intensity in the ranges of $1560\text{--}1570$, $1400\text{--}1460$, and $1000\text{--}1030 \text{ cm}^{-1}$ allows for the complete separation of HBV- and HCV-infected serums from normal serums. For HCV cases, the viral structure and non-structural proteins are created by the cleavage of the polyprotein precursors of the host cell signal peptidase and the viral proteases [54]. The structural core protein is derived from the amino terminus of the polyprotein and may form vi-

rion nucleocapsid because they are highly basic and similar in many biological properties. The nucleotide and predicted amino acid sequences of the core protein gene are highly conserved in different HCV isolates. Earlier investigations have shown that truncated proteins covering the core protein (HCVc-120) are assembled into nucleocapsid-like particles without the presence of nucleic acids with similar morphological and size distributions in the serum of hepatocytes from HCV-infected patients. This explains the abnormal levels of liver enzymes, such as alanine aminotransferase (ALT), and the active division of viruses (HBeAg-positive or high-level HBV DNA) in the human blood and liver biopsies, which show moderate disease activities in hepatitis-B cases. The C virus confirms the presence of positive HCV RNA.

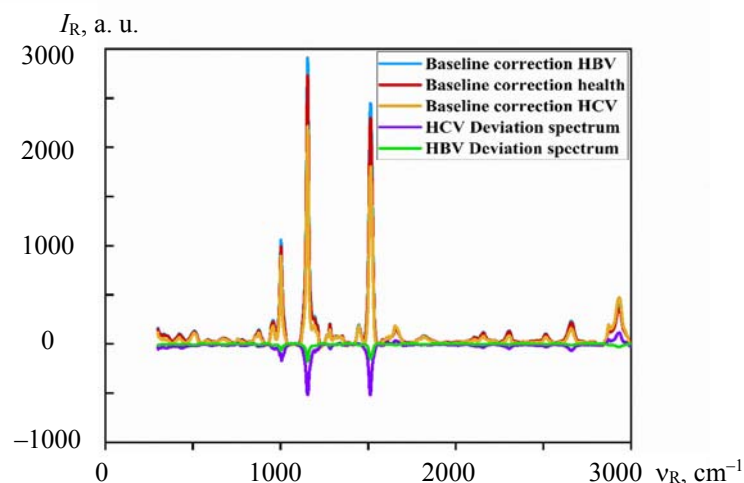


Fig. 2. Raman intensity vs. frequency shift for HBV patients, HCV patients, and healthy subjects.

Table 3 shows the performance of this algorithm for different numbers of PCs. It is evident that the specificity reaches 99.98% when the PC count is 200, and the highest precision of 99.89% is achieved when the PC count is 400. Further, the highest sensitivity and accuracy of 94.54 and 94.92%, respectively, are attained when the PC count is 512, i.e., the best performance is achieved at this PC count. Hence, we set the number of PCs as 512 in subsequent experiments (Table 4).

TABLE 3. Classification Performance Metrics for Different Numbers of PCs

PC, count	Specificity, %	Precision, %	Sensitivity, %	Accuracy, %
60	85.56	71.74	89.19	89.57
100	98.30	95.39	83.78	93.61
200	99.98	98.55	90.54	94.38
300	98.92	97.06	89.19	93.23
400	98.23	99.89	90.54	94.00
500	99.41	99.25	91.89	94.13
512	97.44	98.89	94.54	94.92
525	98.65	99.42	93.66	94.23
550	98.42	99.02	92.34	93.58
600	99.41	98.53	90.54	93.36

TABLE 4. Comparison between the Performance Metrics before and after the Normalization

Normalization Period	Specificity, %	Precision, %	Sensitivity, %	Accuracy, %
Before normalization	96.97	93.06	90.54	90.55
After normalization	97.44	98.89	94.54	94.92

Mean normalization. To prove the effectiveness of mean normalization, we performed two sets of experiments. As shown in Table 3, the utilization of mean normalization increases the specificity by 0.47%, precision by 5.83%, sensitivity by 4.00%, and accuracy by 4.37%.

Comparative experiments. To verify the efficacy of the proposed network, we conducted three comparative experiments using: different layers in our network; commonly used algorithms in deep learning; commonly used algorithms in machine learning.

Feature scales can be extracted through various methods. Too few or too many feature scales can easily lead to underfitting or overfitting, respectively. To compare the stability and goodness of fit for different classifiers, we performed comparative experiments on multi-scale fusion of different layers. Here, all the compared algorithms were implemented in the PyCharm environment.

The experimental results in Table 5 show that with two extraction and fusion layers, the specificity, precision, sensitivity, and accuracy are 97.44, 98.89, 94.54, and 94.92%, respectively. With one extraction and fusion layer, the experimental results show underfitting, and the accuracy is reduced by 2.43% compared to the two-layer scheme. With three extraction and fusion layers, the experimental results show overfitting, and the accuracy is reduced by 3.22% compared to the two-layer scheme. Compared to the one-layer and three-layer schemes, the extraction and fusion layers provide better performance in a two-layer scheme. Overall, these results indicate that the two-layer multi-scale fusion scheme exhibits better stability than the one-layer multi-scale fusion scheme and has better fit than the three-layer multi-scale fusion scheme.

TABLE 5. Experimental Results for Multi-Scale Three-Class Hepatitis Classification Problem with Different Layer Numbers

Number of layers (feature extraction and fusion layers)	Specificity, %	Precision, %	Sensitivity, %	Accuracy, %
One layer	80.23	63.44	79.73	92.55
Two layers	97.44	98.89	94.54	94.92
Three layers	99.39	98.53	90.54	91.70

For comparison of deep learning architectures, we conducted three sets of experiments. As shown in Table 6, we also tried CNNs, whose fitness is inferior to that of MSCNN. In this experiment, although the specificity of CNN reaches 99.18%, its accuracy is only 68.56%. Indeed, single-convolution kernels tend to remove more spectral information, while multi-scale deep schemes improve the accuracy of spectral classification. Features of the input spectral intensity are extracted using convolution kernels of different scales to enhance the robustness of algorithm to interfering factors.

TABLE 6. Performance Comparison among the Three Deep Learning Architectures

Network structure	Specificity, %	Precision, %	Sensitivity, %	Accuracy, %
CNN	99.18	89.47	67.56	68.56
MSCNN	97.44	98.89	94.54	94.92
LSTM	92.03	83.33	87.84	88.89

To compare MSCNN with different machine learning methods, we conducted eight sets of experiments. Here, we also used conventional classifiers including those of naive Bayes, KNN, logistic regression, random forests, decision trees, SVMs, GBDT along with MSCNN for the three-class hepatitis classification problem. The experimental results indicate that conventional classifiers are not adequate for this classification problem, but few of these classifiers still have some advantages. The experimental results are shown in Table 7. Here, for the naive Bayes classifier, it is assumed that the input features are independent of each other. However, this assumption is not practically true because the number of spectral features is relatively large and therefore the correlation between these features is large. Consequently, for the naive Bayes classifier, the accuracy is quite low, and the specificity is 7.97% lower than that of MSCNN.

Generally, the KNN classifier does not perform well for multi-dimensional data including Raman spectral data. High dimensionality of data leads to the ‘curse of dimensionality,’ and it increases the distance between two apparently similar spectral features. Consequently, the sensitivity of KNN is much higher than that of the naïve Bayes classifier, but it is still far lower than that of MSCNN.

TABLE 7. Performance Comparison between MSCNN and other Traditional Classifiers

Algorithm	Recall, %	F1-score, %	Support	Specificity, %	Precision, %	Sensitivity, %	Accuracy, %
Naive Bayes	44.12	37.43	131	89.47	50.00	16.22	43.51
KNN	60.53	60.53	131	80.82	57.58	51.35	59.54
LR	71.87	68.56	131	69.51	56.24	86.49	69.94
RF	65.38	64.94	131	100	100	24.32	64.89
DT	69.55	69.55	131	82.28	57.58	51.14	64.12
SVM	38.73	21.29	131	74.44	58.18	86.47	75.57
GBDT	82.05	82.05	131	91.86	80.00	75.68	82.45
MSCNN	93.00	93.00	131	97.44	98.89	94.54	94.92

Logistic regression cannot handle a large number of multi-class spectral features efficiently. While the sensitivity for this classifier is relatively high, it is still lower than that of MSCNN by 8.05%.

The decision tree classifier ignores the correlation between the Raman spectral features, which leads to low performance metrics and high overfitting.

Overfitting problems can be reduced using random forests, which are equivalent to multiple decision trees whose predictions are pooled to obtain an ensemble prediction. For the random forest classifier, the specificity and precision are perfectly 100%, but the sensitivity is extremely low.

While random forests suffer from a low sensitivity, the GBDT classifier has a relatively high sensitivity, and its accuracy (82.45%) comes second after MSCNN. However, due to the dependence between the weak classifiers of GBDT, it is difficult to train these classifiers in parallel.

SVM algorithm was originally designed for binary classification problems. For multiple classes, it is necessary to construct a suitable multi-class classifier. We classify the hepatitis-B and non-hepatitis-B samples, and then classify the non-hepatitis-C and hepatitis-C samples into two categories. This disturbs the sample balance, and the results are generally inferior for the rare category, namely the hepatitis category. In general, traditional machine learning algorithms are not very effective for the three-category classification problem.

Model evaluation. The classification results of MSCNN are specificity 0.974 ± 0.011 , precision 0.989 ± 0.015 , sensitivity 0.945 ± 0.008 , accuracy 0.949 ± 0.010 . It is clear that the MSCNN model outperforms the other models in terms of accuracy and sensitivity. Although the specificity of CNN reaches 99.18%, a CNN with a single convolution kernel misses more details during information encoding. While random forests show 100% specificity and precision, their accuracy and sensitivity are quite inferior to those of MSCNN.

For a better experimental evaluation, GPU acceleration was used with an NVIDIA 1060 graphics card. In this experiment, we chose a setup where dimensionality reduction was firstly applied with 512 PCs. Then, the features were normalized by the mean value, and finally the MSCNN architecture was used for classification. Averaging the performance over ten runs results in an accuracy of 98.89%, sensitivity of 97.44%, specificity of 94.54%, and accuracy of 94.92%. The confusion matrix

Confusion Matrix		Predicted Class		
		0	1	2
True Class	0	70	1	3
	1	1	95	4
	2	1	4	82

shows the classification outputs for one run.

Conclusions. We demonstrated the effectiveness of using deep learning MSCNN model and serum Raman spectroscopy to screen the HBV-infected and HCV-infected patients and differentiate them from healthy subjects. The baseline correction of spectral data revealed interesting differences between the three classes in terms of biomolecular structure and the relative contents of lipids, cholesterol, proteins, cysteine, etc. The classification performance analysis of MSCNN revealed that the pull Mann's spectrum of serum can be efficiently used to diagnose potential HBV-infected and HCV-infected cases. Owing to the rapid advancement and utilization of Raman spectroscopy, our work is expected to provide accurate and rapid guid-

ance for the treatment of HBV-infected and HCV-infected patients in the near future.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (NSFC) (No. 61765014); Reserve Talents Project of National High-level Personnel of Special Support Program (QN2016YX0324); Urumqi Science and Technology Project (No. P161310002 and Y161010025); and Reserve Talents Project of National High-level Personnel of Special Support Program (Xinjiang [2014]22).

Whether there is a COI or not, such a separate COI page should be established immediately before the References stating the authors take responsibility for any conflict of interest.

REFERENCES

1. H. Nawaz, N. Rashid, M. Saleem, M. A. Hanif, M. I. Majeed, I. Amin, M. Iqbal, M. Rahman, O. Ibrahim, S. M. Baig, M. Ahmed, F. Bonnier, H. J. Byrne, *J. Raman Spectrosc.*, **48**, 697–704 (2017).
2. H. Momose, S. Matsuoka, A. Murayama, N. Yamada, K. Okuma, E. Ikebe, Y. Hoshi, M. Muramatsu, T. Wakita, K. Toyota, T. Kato, I. Hamaguchi, *J. Clin. Virol.*, **105**, 97–102 (2018).
3. J. Liang, C. Lv, M. Chen, M. Xu, C. Zhao, Y. Yang, J. Wang, D. Zhu, J. Gao, R. Rong, T. Zhu, M. Yu, *J. Diabetes*, **11**, 370–378 (2019).
4. Q. Meng, C. Wong, A. Rangachari, S. Tamatsukuri, M. Sasaki, E. Fiss, L. Cheng, T. Ramankutty, D. Clarke, H. Yawata, Y. Sakakura, T. Hirose, C. Impraim, *J. Clin. Microbiol.*, **39**, 2937–2945 (2001).
5. X. Zheng, G. Lv, Y. Zhang, X. Lv, Z. Gao, J. Tang, J. Mo, *Spectrochim. Acta A*, **215**, 244–248 (2019).
6. C. S. Ho, N. Jean, C. A. Hogan, L. Blackmon, S. S. Jeffrey, M. Holodniy, N. Banaei, A. A. E. Saleh, S. Ermon, J. Dionne, *Nat. Commun.*, **10**, 8 (2019).
7. T. Mahmood, H. Nawaz, A. Ditta, M. I. Majeed, M. A. Hanif, N. Rashid, H. N. Bhatti, H. F. Nargis, M. Saleem, F. Bonnier, H. J. Byrne, *Spectrochim. Acta A*, **200**, 136–142 (2018).
8. S. Feng, R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen, H. Zeng, *Biosens. Bioelectron.*, **25**, 2414–2419 (2010).
9. S.-X. Li, Q.-Y. Zeng, L.-F. Li, Y.-J. Zhang, M.-M. Wan, Z.-M. Liu, H.-L. Xiong, Z.-Y. Guo, S.-H. Liu, *J. Biomed. Opt.*, **18**, 027008 (2013).
10. J. Wang, D. Lin, J. Lin, Y. Yu, Z. Huang, Y. Chen, J. Lin, S. Feng, B. Li, N. Liu, R. Chen, *J. Biomed. Opt.*, **19**, 087003 (2014).
11. S. M. Cohen, P. Davitkov, *Liver Disease: A Clinical Casebook*, Springer (2018).
12. Y. LeCun, Y. Bengio, G. Hinton, *Nature*, **521**, 436–444 (2015).
13. J. Acquarelli, T. van Laarhoven, J. Gerretzen, T. N. Tran, L. M. C. Buydens, E. Marchiori, *Anal. Chim. Acta*, **954**, 22–31 (2017).
14. J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, S. J. Gibson, *Analyst.*, **142**, 4067–4074 (2017).
15. C. Cui, T. Fearn, *Chemom. Intell. Lab. Syst.*, **182**, 9–20 (2018).
16. S. L. Neal, *Appl. Spectrosc.*, **72**, 102–113 (2018).
17. S. Malek, F. Melgani, Y. Bazi, *J. Chemom.*, **32**, e2977 (2018).
18. N. H. Tran, X. Zhang, L. Xin, B. Shan, M. Li, *P Natl. Acad. Sci. USA*, **114**, 8247–8252 (2017).
19. X.-X. Zhou, W.-F. Zeng, H. Chi, C. Luo, C. Liu, J. Zhan, S.-M. He, Z. Zhang, *Anal. Chem.*, **89**, 12690–12697 (2017).
20. S. Wang, S. Fei, Z. Wang, Y. Li, J. Xu, F. Zhao, X. Gao, *Bioinformatics*, **35**, 691–693 (2019).
21. P. Inglese, J. S. McKenzie, A. Mroz, J. Kinross, K. Veselkov, E. Holmes, Z. Takats, J. K. Nicholson, R. C. Glen, *Chem Sci.*, **8**, 3500–3511 (2017).
22. M. Wen, Z. Zhang, S. Niu, H. Sha, R. Yang, Y. Yun, H. Lu, *J. Proteome Res.*, **16**, 1401–1409 (2017).
23. M. Wen, P. Cong, Z. Zhang, H. Lu, T. Li, *Bioinformatics*, **34**, 3781–3787 (2018).
24. D. Jiang, S. Malla, Y.-J. Fu, D. Choudhary, J. F. Rusling, *Anal. Chem.*, **89**, 12872–12879 (2017).
25. N. Divakar, R. Venkatesh Babu, *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops*, 80–87 (2017).
26. S. Li, G. Chen, Y. Zhang, Z. Guo, Z. Liu, J. Xu, X. Li, L. Lin, *Opt. Express*, **22**, 25895–25908 (2014).
27. R. Sumazaki, M. Motz, H. Wolf, J. Heinig, W. Jilg, F. Deinhardt, *J. Med. Virol.*, **27**, 304–308 (1989).
28. J. Zhao, H. Lui, D. I. McLean, H. Zeng, *Appl. Spectrosc.*, **61**, 1225–1232 (2007).
29. X. Li, T. Yang, S. Li, D. Wang, Y. Song, S. Zhang, *Laser Phys.*, **26**, 035702 (2016).
30. A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk, M. Baranska, *J. Raman Spectrosc.*, **44**, 1061–1076 (2013).
31. V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, doi **10**, 978–971 (1995).

-
32. I. Düntsch, G. Gediga, *Conf. Ser.*, IOP Publishing., **1229**, 012055 (2019).
 33. G. P. J. Yang, V. Rao, J. Sohl-Dickstein, S. S. Schoenholz, arXiv preprint arXiv:1902.08129 (2019).
 34. G. Yang, J. Pennington, V. Rao, J. Sohl-Dickstein, S. S. Schoenholz, arXiv preprint arXiv:1902.08129 (2019).
 35. Q. Luo, H. Ma, Y. Wang, L. Tang, R. Xiong, *Neurocomputing*, **378**, 364–374 (2020).
 36. X.-X. Zhou, H. Chi, C. Luo, C. Liu, J. Zhan, et al. *Anal. Chem.*, **89**, 12690–12697 (2017).
 37. Z. Movasaghi, S. Rehman, I. U. Rehman, *Appl. Spectrosc. Rev.*, **42**, 493–541 (2007).
 38. L. Seballos, J. Z. Zhang, R. Sutphen, *Anal. Bioanal. Chem.*, **383**, 763–767 (2005).
 39. K. Maquelin, C. Kirschner, L.-P. Choo-Smith, N. van den Braak, D. Endtz, H. P. Naumann, G. Puppels, *J. Microbiol. Meth.*, **51**, 255–271 (2002).
 40. C.-C. Lin, W.-H. Liu, Z.-H. Wang, M.-C. Yin, *Eur. J. Nutr.*, **50**, 499–506 (2011).
 41. N. Kato, O. Yokosuka, M. Omata, K. Hosoda, M. Ohto, *J. Clin. Invest.*, **86**, 1764–1767 (1990).
 42. N. Stone, C. Kendall, J. Smith, P. Crow, H. Barr, *Faraday Discuss.*, **126**, 141–157 (2004).
 43. S. Anwar, S. Firdous, *Laser Phys. Lett.*, **12**, 076001 (2015).
 44. J. Perla-Kaján, H. Jakubowski, *Amino Acids*, **43**, 1405–1417 (2012).
 45. G. Shetty, C. Kendall, N. Shepherd, N. Stone, H. Barr, *Brit. J. Cancer.*, **94**, 1460–1464 (2006).
 46. A. A. Raouf, H. M. El-Sebaey, A. K. Abd El-Hamead, A. Y. El-Fert, Y. E. El-Gendy, *Menoufia Med. J.*, **29**, 895 (2016).
 47. T. Hevonoja, M. O. Pentikäinen, M. T. Hyvönen, P. T. Kovanen, M. Ala-Korpela, *BBA-Mol Cell Biol L.*, **1488**, 189–210 (2000).
 48. C. M. Bremer, C. Bung, N. Kott, M. Hardt, D. Glebe, *Cell Microbiol.*, **11**, 249–260 (2009).
 49. Y.-J. Li, P. Zhu, Y. Liang, W.-G. Yin, J.-H. Xiao, *World J. Gastroenter.*, **19**, 2262 (2013).
 50. S. Khan, R. Ullah, A. Khan, R. Ashraf, H. Ali, M. Bilal, M. Saleem, *Photodiagn. Photodyn.*, **23**, 89–93 (2018).
 51. K. Naseer, M. Saleem, S. Ali, B. Mirza, J. Qazi, *Spectrochim. Acta A*, **222**, 117181 (2019).
 52. A. Rodríguez-Casado, M. Molina, P. Carmona, *Appl. Spectrosc.*, **61**, 1219–1224 (2007).
 53. A. Ditta, H. Nawaz, T. Mahmood, M. Majeed, M. Tahir, N. Rashid, M. Muddassar, A. Al-Saadi, H. Byrne, *Spectrochim. Acta A*, **221**, 117173 (2019).
 54. Y. Lu, Y. Lin, Z. Zheng, X. Tang, J. Lin, X. Liu, M. Liu, G. Chen, S. Qiu, T. Zhou, *Biomed. Opt. Express*, **9**, 4755–4766 (2018).