RESEARCH ARTICLE

Deep learning-based ultra-fast identification of Raman spectra with low signal-to-noise ratio

Kunxiang Liu[1,2] 🔘　|　**Fuyuan Chen**[1,2]　|　**Lindong Shang**[1,2]　|　**Yuntong Wang**[1,2]　|　**Hao Peng**[1,2]　|　**Bo Liu**[1,2] 🔘　|　**Bei Li**[1,2]

[1]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, PR China

[2]University of Chinese Academy of Sciences, Beijing, PR China

Correspondence

Bei Li, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, PR China.
Email: beili@ciomp.ac.cn

Bo Liu, University of Chinese Academy of Sciences, Beijing 100049, PR China.
Email: liubo182@mails.ucas.ac.cn

Abstract

Ensuring the correct use of cell lines is crucial to obtaining reliable experimental results and avoiding unnecessary waste of resources. Raman spectroscopy has been confirmed to be able to identify cell lines, but the collection time is usually 10–30 s. In this study, we acquired Raman spectra of five cell lines with integration times of 0.1 and 8 s, respectively, and the average accuracy of using long–short memory neural network to identify the spectra of 0.1 s was 95%, and the average accuracy of identifying the spectra of 8 s was 99.8%. At the same time, we performed data enhancement of 0.1 s spectral data by real-valued non-volume preserving method, and the recognition average accuracy of long–short memory neural networks recognition of the enhanced spectral data was improved to 96.2%. With this method, we shorten the acquisition time of Raman spectra to 1/80 of the original one, which greatly improves the efficiency of cell identification.

KEYWORDS

cell line identification, data enhancement, deep learning, Raman spectroscopy
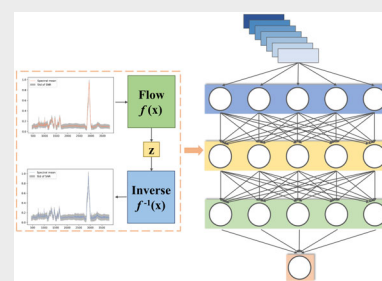
1 | INTRODUCTION

Cell lines are utilized frequently in life science and clinical research as a common material in investigations like cytokine activity tests and the screening of anti-tumor drugs [1, 2]. Cell line abuse brought on by contaminated cell lines can directly affect the precision and dependability of scientific research and lead to time and resource waste. The National Institutes of Health (NIH) and other organizations have recently requested that researchers identify cells before performing studies [3]. The currently prevalent methods for identifying cell lines, such as short tandem repeat (STR) analysis, often require complex operations such as staining or culturing of cells [4]. Given

the successful applications of Raman spectroscopy in various fields, such as microbial identification, species identification, food identification, microplastic identification, drug analysis, and tumor diagnosis [5–12], we utilized Raman spectroscopy and a self-built SL-Raman model to accurately distinguish between normal and gastric cancer cell lines in our previous study [13].

In our previous study, we set the acquisition condition for Raman spectra of animal cells to 8 s, and the acquisition time of most Raman spectra of animal cells is about 10–30 s. However, this acquisition time is considerably longer than the 1–5 s acquisition time reported in other studies of microbial Raman spectra. For cell or tissue samples, it usually takes hours to obtain spectral

signals over a large area or to obtain high-resolution Raman spectroscopic imaging, which is fatal for the application of high-throughput Raman spectroscopy. To address this issue, we aim to improve the efficiency of spectral acquisition and shorten the identification time of cell lines using the real-valued non-volume preserving (RealNVP) algorithm for spectral data enhancement and long-short memory neural networks (LSTM). Specifically, we plan to utilize these techniques to achieve an acquisition time of 0.1 s for identifying Raman spectra of cell lines. By leveraging these approaches, we hope to significantly reduce the acquisition time of Raman spectra in cell line studies and enhance the efficiency of spectral acquisition without compromising the average accuracy of cell line identification.

The mainstream generative models commonly used are variational self-encoder (VAE), generative adversarial network (GAN), and flow-based generative models. VAE stores potential attributes as probability distributions, but its method of recovering data distribution and calculating loss function can result in a vague generated output [14]. GAN is a deep learning-based generative model that can generate new content, but it only distinguishes between "real" and "fake" images, which can lead to generated images without actual objects but with similar-looking styles [15, 16]. In contrast, the flow-based generative model is constructed by a sequence of reversible transformations and learns the data distribution explicitly. Therefore, the loss function is only negative log-likelihood. In this study, we employ the upgraded stream-based generative model, RealNVP, to generate Raman spectral data with an integration time of 0.1 s.

Machine learning and deep learning methods are commonly utilized in the study of Raman spectral data identification. However, traditional machine learning methods such as K-Nearest Neighbor (KNN) and linear discriminant analysis (LDA) can reduce the predictive power of the models when the sample source, spectral acquisition conditions, and spectral instruments change [11]. Convolutional neural networks (CNN) such as Resnet and visual geometry group (VGG) are adept at extracting features from data but are better suited for processing image data. Additionally, when there are many spectral categories to be classified, increasing the depth of the CNN network can improve classification average accuracy, but this requires more computation and time [15, 17, 18]. Recurrent neural networks (RNN) were developed to consider the correlation of data before and after processing sequential data. The primary method is to add the data preceding (and/or following) the data as a weight to the original inference. Compared to CNN, RNN is more suitable for processing one-dimensional Raman spectral data [19]. However, RNN has the issue of gradient explosion or long-term dependence. To address these issues, we utilize the LSTM approach. LSTM adds a forgetting mechanism to RNN, making it possible to selectively retain or forget certain data from the previous period. It also avoids the problem of gradient explosion by using an additive method instead of the multiplicative iteration used in RNN [14].

To enhance the efficiency of cell line identification and shorten the time of Raman spectrum acquisition, we employed RealNVP for data enhancement of Raman spectra with an integration time of 0.1 s and LSTM for Raman spectrum identification of cell lines with low signal-to-noise ratio. We initially attempted to identify spectral data with low signal-to-noise ratios using KNN and LDA for the 0.1 and 8 s Raman spectra, respectively, but found it to be challenging. We subsequently utilized LSTM to identify the 0.1 and 8 s Raman spectral data and achieved high recognition average accuracy for both. Furthermore, the average accuracy of LSTM recognizing the 0.1 s Raman spectra was improved after data augmentation. By combining data augmentation and LSTM, we reduced the acquisition time of Raman spectra for cell line identification to 1/80 of the original time. This approach significantly improves the efficiency of spectral acquisition without compromising the average accuracy of cell line identification and has promising applications in the field of life science and clinical research.

2 | MATERIALS AND METHODS

2.1 | Sample preparation

In this study, we utilized one normal cell line (GES-1) and four gastric cancer cell lines (AGS, HGC-27, MKN-45, and MKN-74). The cells were preserved in a $-80°C$ freezer and defrosted in a $37°C$ water bath before sample preparation. AGS cells were cultured in DMEM/F12 medium supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin, while the other cell lines were cultured in RPMI-1640 medium with 10% fetal bovine serum and 1% penicillin–streptomycin. All cell lines were cultured in a 5% CO_2 humidified incubator at $37°C$.

After 48 h of culturing in a cell culture flask, the cells were removed with 0.25% trypsin–EDTA and washed three times with phosphate-buffered saline. The collected suspended cells were then fixed in sterilized deionized water with 4% paraformaldehyde, diluted to 10 000 per milliliter using deionized water, immobilized on a slide, and air-dried for Raman spectra measurements. The Raman detection sample for each cell line was approximately 7.5 µL (2.5 µL/drop ×3). We used a glass slide coated with a 25 nm thick aluminum film in the experiment. All reagents used in this process were purchased from Gibco Company, USA.

2.2 | Raman measurements

In this study, we utilized a Raman spectroscopy system from Hooke Instruments, Changchun, China (R300 (objective: Olympus, 100×, NA = 0.8), with a laser wavelength of 532 nm) to collect Raman spectra of six gastric cancer cell lines. Due to the difference between animal cells and the laser spot size, the Raman spectra collected at a single location were not representative of the biochemical composition information of the entire cells. Therefore, we randomly collected four to five spectra at different locations of each cancer cell to ensure the representativeness of the biochemical composition information. The Raman spectra acquisition conditions were set at 600 grating, a laser power of 9 mW, and integration times of 0.1 and 8 s. To avoid irreversible damage to the cells caused by the integration time of 8 s, we first acquired the spectra of 0.1 s, followed by the spectra of 8 s at the same position. Hence, the material information corresponding to the two batches of Raman spectra with different integration times was the same.

Finally, we collected a total of 1000 (500 for 0.1 s and 500 for 8 s) Raman spectra of the normal cell line GES-1, 1010 Raman spectra of gastric cancer cell line AGS, 1002 Raman spectra of gastric cancer cell line HGC-27, 1060 Raman spectra of gastric cancer cell line MKN-45, and 1008 Raman spectra of gastric cancer cell line MKN-74.

2.3 | Data preprocessing

To address the minor disparities in the x-axis across the various spectra, we utilized a three-fold spline interpolation technique to process the data. The preprocessing of Raman spectra obtained under distinct integration time conditions involved the removal of cosmic rays, baseline correction, and normalization. Adaptive iterative re-weighted penalized least squares (airPLS) algorithm was employed for baseline correction, and min-max normalization was used for normalization [20, 21]. The preprocessed spectral data ranged from 400 to 3820 cm^{-1}, and the number of spectral data features was 1326. The process of spectral preprocessing was implemented using the Python programming language.

3 | RESULTS AND DISCUSSION

3.1 | KNN and LDA based identification of gastric cancer cell lines

We have acquired a total of 5080 Raman spectra (2540 for 0.1 s and 2540 for 8 s) of one normal gastric epithelial cell line and four gastric cancer cell lines, as presented in Table 1. Figure 1 displays the mean spectra of Raman spectra obtained at different integration times, while Figure 1B represents the signal-to-noise box plots of these spectra. It is evident that the noise of the Raman spectral signal is slightly higher for 0.1 s, and the signal-to-noise ratio ranges from 0 to 40 for the 0.1 s data and 0 to 400 for the 8 s data. The Raman spectra of the cells obtained at an integration time of 8 s are of superior quality.

We employed two traditional machine learning techniques, KNN and LDA, to classify the two datasets. Figure 2A,B illustrates the confusion matrices of KNN for classifying 0.1 and 8 s gastric cancer cell lines, respectively, where the number of spectra in the test set is 100 per class. KNN achieved an average accuracy of 59% for identifying 0.1 s Raman spectral data and 94.8% for identifying 8 s Raman spectral data. Figure 2C,D presents the confusion matrix of LDA for classifying 0.1 and 8 s gastric cancer cell lines, respectively. The average accuracy of LDA for identifying 0.1 s Raman spectral data was 88.6%, and for 8 s Raman spectral data, it was 99%. KNN faced difficulty in recognizing spectral data with a poor signal-to-noise ratio. To improve the utilization of 0.1 s data for recognizing 8 s data, we may need to explore alternative methods.

3.2 | RealNVP-based Raman spectral data generation

To be able to better utilize the poor-quality data of 0.1 s, we intend to first perform data enhancement of Raman spectra using the Flow data generation algorithm (Figure 3A), and then identify the spectral data using LSTM (Figure 3B).

RealNVP is an advanced version of the NICE model, which was one of the pioneering Flow models, proposed by Dinh [22]. RealNVP generalizes the coupled layers and successfully incorporates convolutional layers in the coupled model, enabling improved handling of image problems. Additionally, it proposes the use of multiscale layers, which can reduce the computational burden and provide a powerful regularization effect, thereby enhancing the quality of generated samples.

After collecting the 0.1 s Raman spectra of cell lines, we preprocessed the data by randomly selecting 100 spectra of each normalized spectral data type as the test set. The remaining Raman spectra were used as the training set for the RealNVP data generation network. The RealNVP network consisted of 8 hidden layers with 256 neurons each, and the network was trained for 1000 epochs using a batch size of 64. The amount of generated

data was set to 1600, and the number of spectral features was 1326.

Figure 4 illustrates a comparison between the generated Raman spectra and the mean spectrum of the original Raman spectra. The red curve represents the average spectrum of the acquired data, while the black curve represents the average spectrum of the algorithm-generated spectral data. It can be observed from the two average spectra that the RealNVP-generated Raman spectral data exhibit slightly more divergence and a slightly weaker signal-to-noise ratio than the acquired real data.

TABLE 1 Number of Raman spectra of cell lines collected under different integration time conditions.

Cell line name	Number of spectra (0.1 s)	Number of spectra (8 s)
GES-1	500	500
AGS	505	505
HGC-27	501	501
MKN-45	530	530
MKN-74	504	504
Total	2540	2540

Figure 5A,B displays 2D and 3D plots of PCA downscaling of Raman spectral data (acquired real spectrum and algorithm-generated spectrum) of cell line GES-1 (real spectrum $n = 400$, generated spectrum $n = 1600$), respectively. We can see that the distribution ranges of the two sets of data are essentially the same. Figure 5C shows the confusion matrix of the two sets of data identified by KNN. This matrix classifies the two sets of data as one class, indicating that the data generated using RealNVP is not significantly different from the original data.

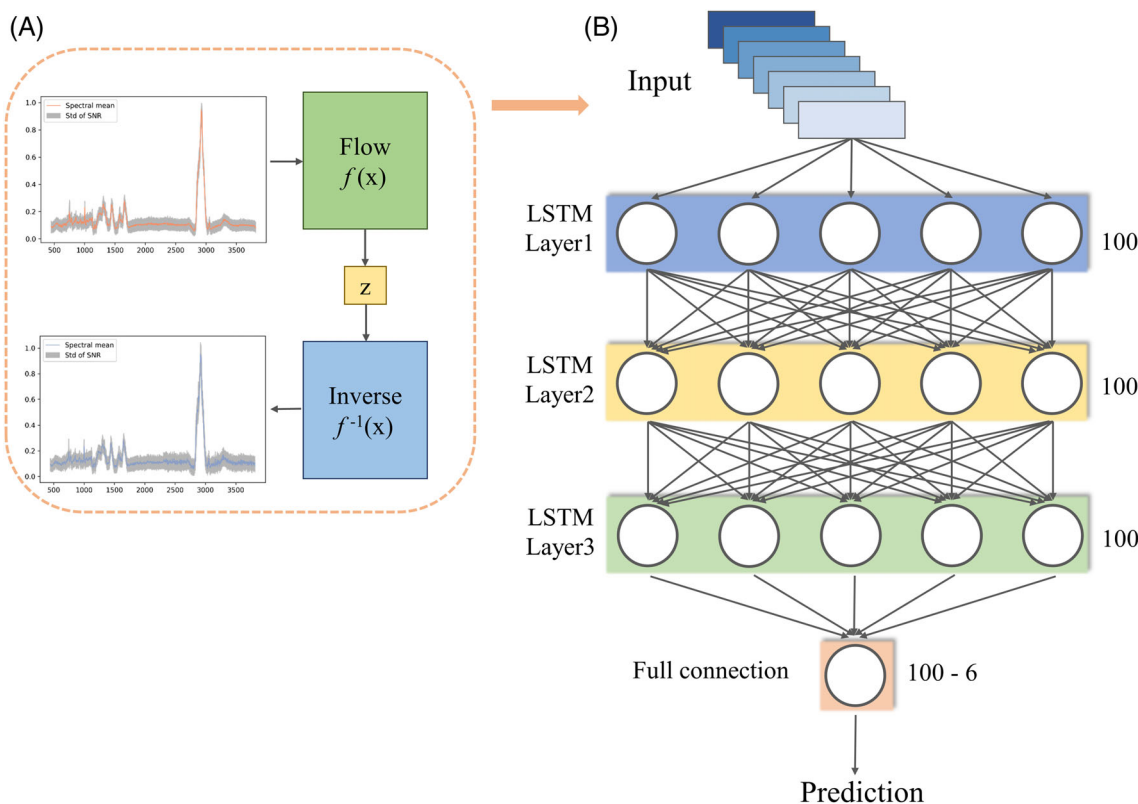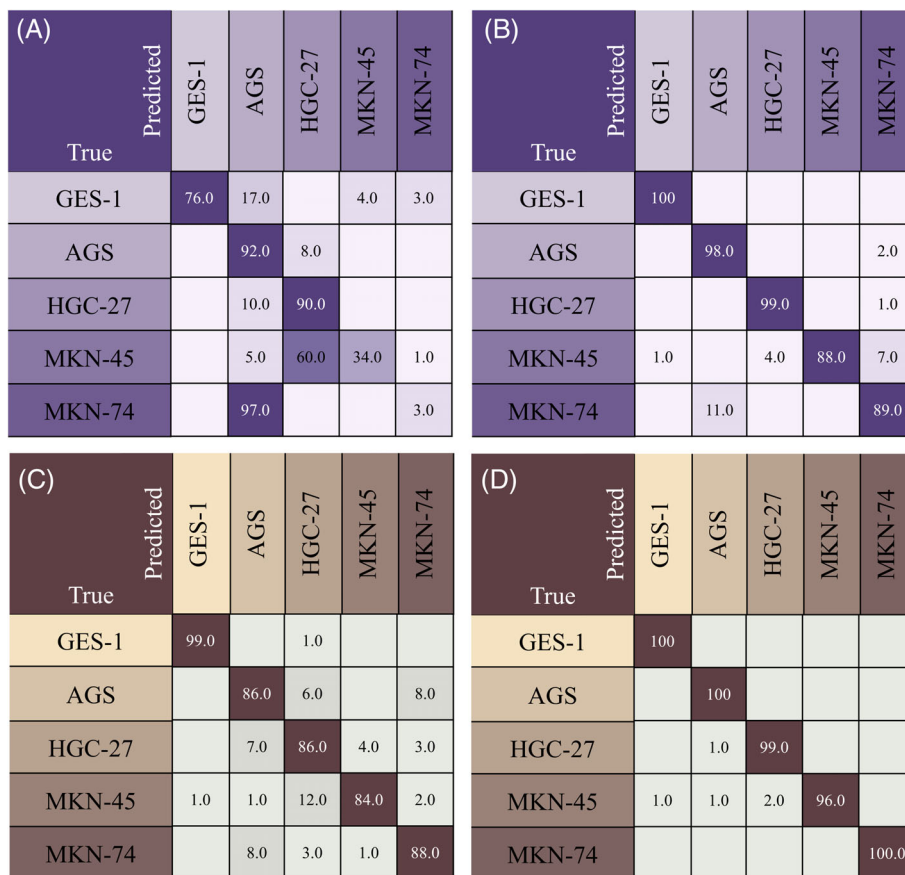## 3.3 | LSTM-based identification of gastric cancer cell lines

LSTM is an improved version of RNN that addresses the issue of long-term dependencies in RNNs. LSTM incorporates a forgetting mechanism into RNN, allowing it to selectively retain or forget certain data from the previous time step. Additionally, it uses addition instead of the multiplicative iteration used in RNN to avoid the problem of gradient explosion [23]. The core idea of LSTM is that the previous memory (data) is not output in its entirety. Instead, when selecting to output the previous data, it adds its own judgment based on three
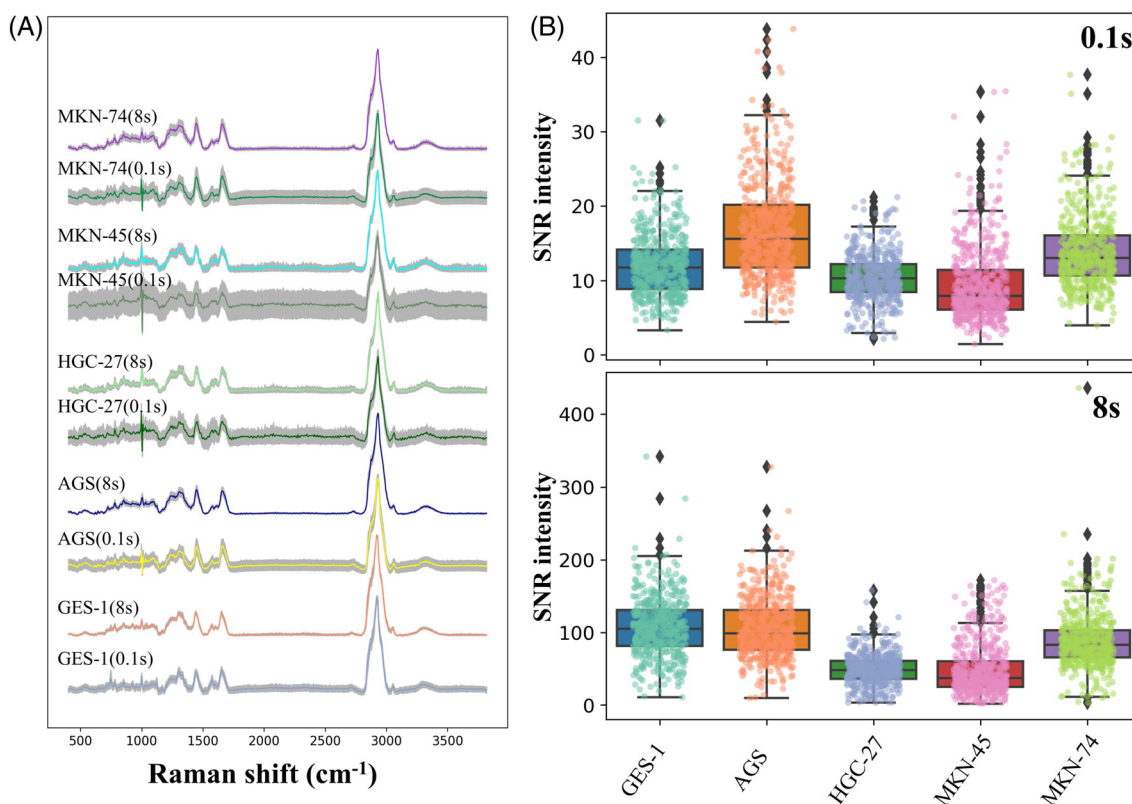

FIGURE 1 Mean Raman spectra and spectral S/N box plots of cell lines. (A) Mean Raman spectra of five cell lines acquired at integration times of 0.1 and 8 s. (B) Raman spectra S/N box plots of five cell lines acquired at integration times of 0.1 and 8 s.

FIGURE 2 Identification results of Raman spectra of gastric cancer cell lines with 0.1 and 8 s identified by KNN and LDA, respectively. (A) Confusion matrix of Raman spectra of gastric cancer cell lines with 0.1 s identified by KNN. (B) Confusion matrix of KNN recognition of 8 s Raman spectra of gastric cancer cell lines. (C) Confusion matrix of LDA recognition 0.1 s Raman spectra of gastric cancer cell lines. (D) Confusion matrix of LDA recognition of 8 s Raman spectra of gastric cancer cell lines.

(A)

True \ Predicted	GES-1	AGS	HGC-27	MKN-45	MKN-74
GES-1	76.0	17.0		4.0	3.0
AGS		92.0	8.0		
HGC-27		10.0	90.0		
MKN-45		5.0	60.0	34.0	1.0
MKN-74		97.0			3.0

(B)

True \ Predicted	GES-1	AGS	HGC-27	MKN-45	MKN-74
GES-1	100				
AGS		98.0			2.0
HGC-27			99.0		1.0
MKN-45	1.0		4.0	88.0	7.0
MKN-74		11.0			89.0

(C)

True \ Predicted	GES-1	AGS	HGC-27	MKN-45	MKN-74
GES-1	99.0		1.0		
AGS		86.0	6.0		8.0
HGC-27		7.0	86.0	4.0	3.0
MKN-45	1.0	1.0	12.0	84.0	2.0
MKN-74		8.0	3.0	1.0	88.0

(D)

True \ Predicted	GES-1	AGS	HGC-27	MKN-45	MKN-74
GES-1	100				
AGS		100			
HGC-27		1.0	99.0		
MKN-45	1.0	1.0	2.0	96.0	
MKN-74					100.0

FIGURE 3 Structure of the algorithm based on RealNVP and LSTM. (A) Structure of the data generation model based on RealNVP. (B) Structure of the cell line Raman spectra identification model based on LSTM.
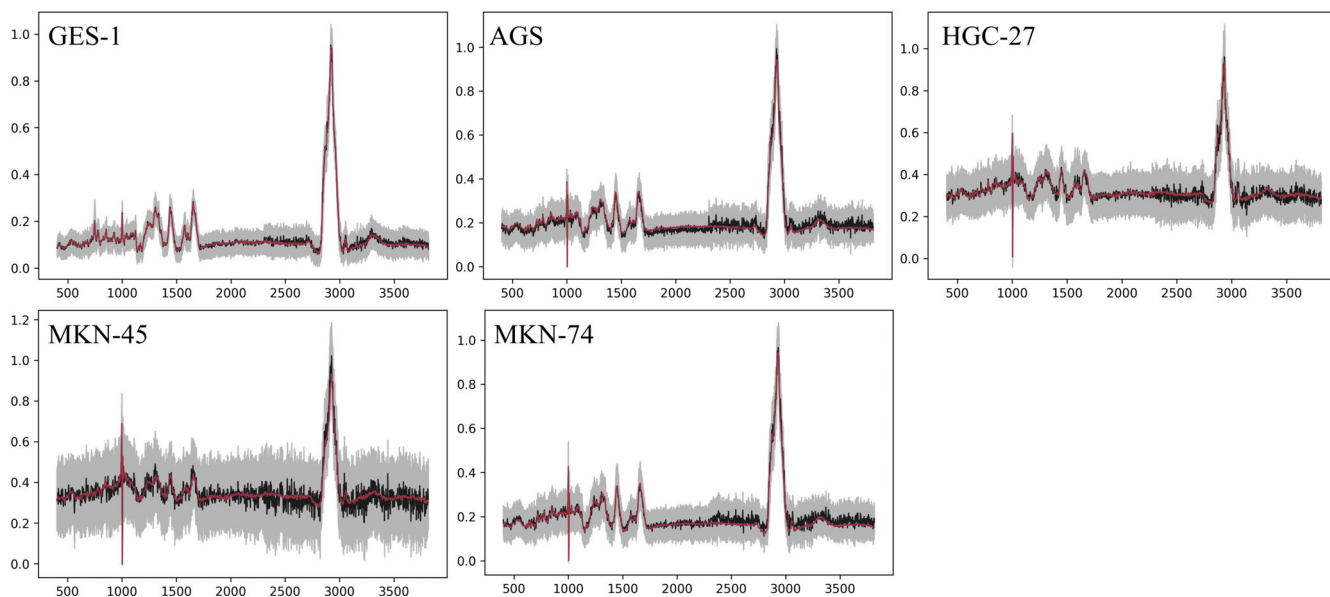
FIGURE 4 Comparison of the real Raman spectra and the Raman spectral data generated by the RealNVP algorithm. The red curve is the real average Raman spectrogram acquired by the spectrometer. The black curve is the average Raman spectrogram generated by the RealNVP algorithm. The gray area is the standard deviation of the spectra generated by the RealNVP algorithm.
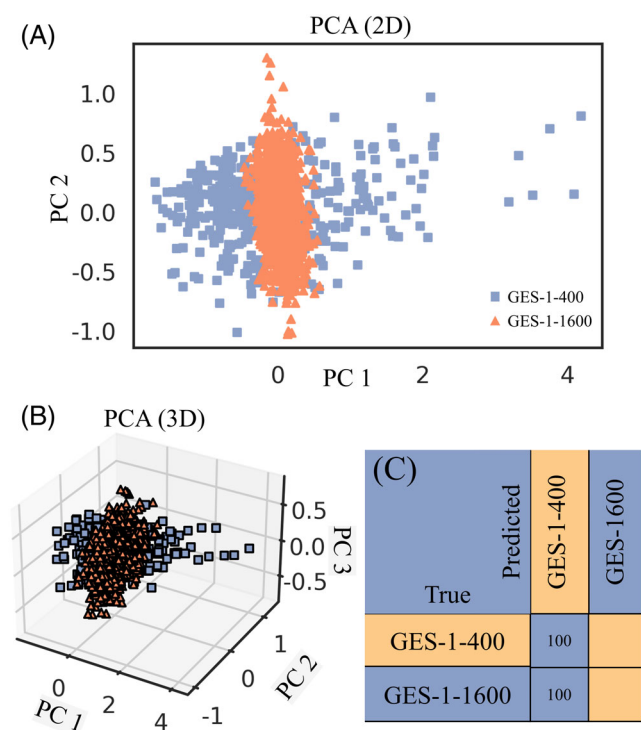

FIGURE 5 Distribution of spectral data and KNN recognition results for PCA downscaling of two data sets of cell line GES-1. (A) Distribution of real spectra ($n = 400$) versus algorithm-generated spectra ($n = 1600$) under two-dimensional PCA. (B) Distribution of real spectra versus algorithm-generated spectra under three-dimensional PCA. (C) KNN recognition of confusion matrix for two data sets.

questions [1]: Should the old content be forgotten based on the new content? [2] How much of the new content needs to be remembered? [3] How much of the remembered content should be taken out to output? By considering these questions, LSTM can effectively capture long-term dependencies in sequential data and has achieved remarkable success in various applications such as natural language processing and speech recognition.

In this study, we utilized a three-layer structured LSTM to identify gastric cancer cell lines. The hidden layer of the LSTM network consisted of 100 nodes, and a five-dimensional output was produced through a fully connected layer using the ReLU activation function. The loss values were calculated using the cross-entropy loss function. Use the cross entropy loss function to calculate the loss value. ADAM optimizer was used to train the network, using the following parameters: the learning rate at 0.0001; exponential decay rates at $\beta_1 = 0.5$ and $\beta_2 = 0.999$.

Figure 6A,B displays the confusion matrix and ROC curves for identifying Raman spectra with an integration time of 8 s using the LSTM network. The training set data contained $n = 2040$ samples, while the test set data contained $n = 500$ samples. The recognition average accuracy of LSTM for the 8 s data was calculated to be 99.6%. Figure 6C,D shows the confusion matrix and ROC curves for identifying Raman spectra with an integration time of 0.1 s using the LSTM network. The training set data contained $n = 2040$ samples, while the test set data
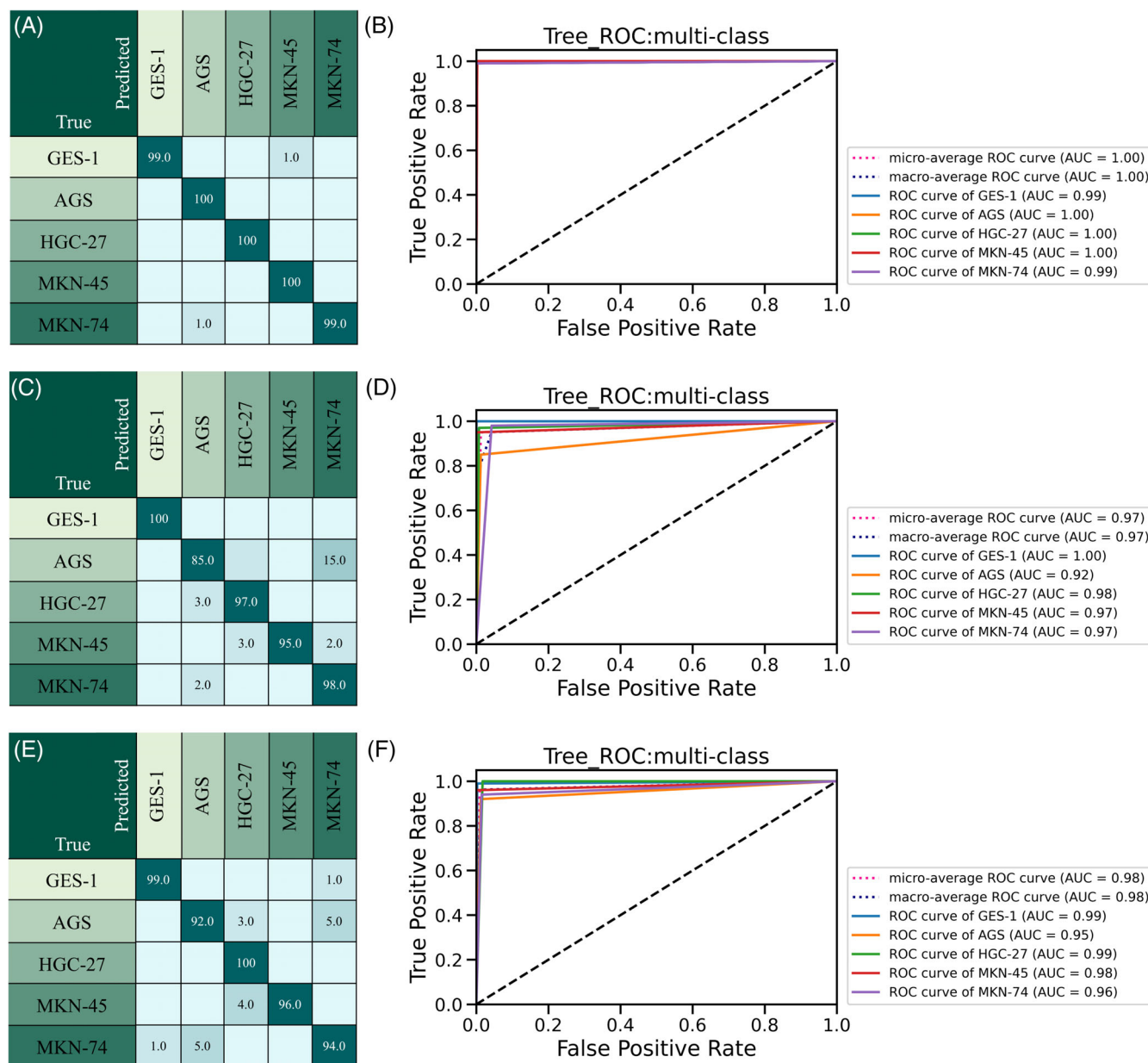
FIGURE 6   Results of LSTM recognition of Raman spectra of gastric cancer cell lines. (A) Confusion matrix of LSTM recognition of Raman spectra of cell lines with 8 s. (B) ROC curve of LSTM recognition of Raman spectra of cell lines with 8 s. (C) Confusion matrix of LSTM recognition of Raman spectra of cell lines with 0.1 s. (D) ROC curve of LSTM recognition of Raman spectra of cell lines with 0.1 s. (E) Confusion matrix of LSTM for the identification of Raman spectra of cell lines after data enhancement. (F) ROC curves of LSTM for the identification of Raman spectra of cell lines after data enhancement.

contained $n = 500$ samples. The recognition average accuracy of LSTM for the 0.1 s data was calculated to be 95%. Finally, Figure 6E,F displays the confusion matrix and ROC curves of Raman spectra with an integration time of 0.1 s after data enhancement using LSTM. The training set data contained $n = 10\,040$ samples, while the test set data contained $n = 500$ samples. The recognition average accuracy of LSTM for the data enhanced with 0.1 s was calculated to be 96.2%.

The results of the study indicate that LSTM outperforms KNN, LDA, and other machine learning methods for recognition of Raman spectral data. The recognition average accuracy of LSTM for Raman spectra with an integration time of 0.1 s is 95%, and the recognition average accuracy after data enhancement using RealNVP is improved to 96.2%. Therefore, RealNVP is effective for enhancing Raman spectral data and improving the classification average accuracy of the classification model.

Although the final recognition average accuracy of the model using the enhanced data is lower than that of the 8 s data, a recognition average accuracy of 96.2% can meet the needs of most application scenarios. Overall, these findings suggest that the combination of RealNVP and LSTM can be a promising approach for enhancing and analyzing Raman spectral data, especially in fast identification scenarios where Raman acquisition times are critical.

4 | CONCLUSIONS

In this study, Raman spectra of one normal gastric epithelial cell line and four gastric cancer cell lines were acquired, including 2540 Raman spectra with an integration time of 0.1 s and 2540 Raman spectra with an integration time of 8 s. The spectral quality of the 0.1 s data was found to be much worse in terms of signal-to-noise ratio.

The results of the study showed that the average accuracy of identifying the Raman spectral data of 0.1 s using KNN was only 59%, while the average accuracy of identifying the Raman spectral data of 8 s was 94.8%. The average accuracy of identifying the Raman spectral data of 0.1 s using LDA was 88.6%, while the average accuracy of identifying the Raman spectral data of 8 s was 99%. These results suggest that the identification and characterization of gastric cancer cell lines can be successfully achieved using machine learning methods with high quality Raman spectral data, but such identification is not possible with spectral data that has an integration time of 0.1 s due to its poor quality.

We employed deep learning LSTM to achieve recognition of Raman spectral data with an integration time of 0.1 s, yielding an impressive recognition average accuracy of 95%. Additionally, for Raman spectral data with an integration time of 8 s, LSTM achieved a recognition average accuracy of 99.8%. Furthermore, we utilized RealNVP to augment the amount of spectral data in the training set with a factor of four, resulting in an increase in the recognition average accuracy of LSTM to 96.2%.

Hence, RealNVP can effectively generate Raman spectral data, reduce the amount of data collection, and enhance the recognition average accuracy of the classification model. Although the recognition average accuracy of RealNVP combined with LSTM is lower than that of LSTM for 8 s spectral data, the recognition average accuracy of 96.2% is sufficient for most application scenarios. Moreover, this method significantly reduces the time required for Raman spectrum acquisition, down to 1/80th of the original time.

In our future research, we aim to standardize the processing of spectral samples, as well as the acquisition conditions and pre-processing parameters of Raman spectroscopy, for the purpose of cell identification. Additionally, we intend to establish a comprehensive Raman spectroscopy database for human cell lines, with the ultimate goal of effectively applying Raman spectroscopy technology to early cell identification.

AUTHOR CONTRIBUTIONS
Kunxiang Liu: Formal analysis, Writing – original draft, Visualization, Implementation. **Fuyuan Chen:** Experiment. **Lindong Shang:** Formal analysis. **Yuntong Wang:** Formal analysis. **Hao Peng:** Writing – original draft. **Bo Liu:** Writing – review & editing. **Bei Li:** Writing – review & editing.

CONFLICT OF INTEREST STATEMENT
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Kunxiang Liu https://orcid.org/0000-0001-9115-3173
Bo Liu https://orcid.org/0000-0002-7856-1989

REFERENCES
[1] B. D. Fera, H. Liu, F. Tian, *Cancer Res.* **2021**, *81*, 1648.
[2] E. Gurgul-Convey, I. Mehmeti, T. Plotz, A. Jorns, S. Lenzen, *Diabetologia* **2016**, *59*, 2125.
[3] R. Barallon, S. R. Bauer, J. Butler, A. Capes-Davis, W. G. Dirks, E. Elmore, M. Furtado, M. C. Kline, A. Kohara, G. V. Los, R. A. F. MacLeod, J. R. W. Masters, M. Nardone, R. M. Nardone, R. W. Nims, P. J. Price, Y. A. Reid, J. Shewale, G. Sykes, A. F. Steuer, D. R. Storts, J. Thomson, Z. Taraporewala, C. Alston-Roberts, L. Kerrigan, *In Vitro Cell Dev. Biol. Anim.* **2010**, *46*, 727.
[4] A. Edwards, A. Civitello, H. A. Hammond, C. T. Caskey, *Am. J. Hum. Genet.* **1991**, *49*, 746.
[5] C. F. Araujo, M. M. Nolasco, A. M. P. Ribeiro, P. J. A. Ribeiro-Claro, *Water Res.* **2018**, *142*, 426.
[6] S. Pahlow, S. Meisel, D. Cialla-May, K. Weber, P. Rosch, J. Popp, *Adv. Drug Delivery Rev.* **2015**, *89*, 105.
[7] A. Saletnik, B. Saletnik, C. Puchalski, *Molecules* **2022**, *27*, 4454.
[8] W. T. Wang, H. Zhang, Y. Yuan, Y. Guo, S. X. He, *AAPS PharmSciTech.* **2018**, *19*, 2921.

[9]   Y. Xu, P. Zhong, A. Jiang, X. Shen, X. Li, Z. Xu, Y. Shen, Y. Sun, H. Lei, *Trends Anal. Chem.* **2020**, *131*, 116017.

[10] L. Shang, L. Xu, Y. Wang, K. Liu, P. Liang, S. Zhou, F. Chen, H. Peng, C. Zhou, Z. Lu, B. Li, *Anal. Methods* **2022**, *14*, 5056.

[11] K. Liu, Q. Zhao, B. Li, X. Zhao, *Front. Bioeng. Biotechnol.* **2022**, *10*, 856591.

[12] B. Liu, K. Liu, X. Qi, W. Zhang, B. Li, *Sci. Rep.* **2023**, *13*, 3240.

[13] K. Liu, B. Liu, Y. Zhang, Q. Wu, M. Zhong, L. Shang, Y. Wang, P. Liang, W. Wang, Q. Zhao, B. Li, *Comput. Struct. Biotechnol. J.* **2023**, *21*, 802.

[14] B. Liu, K. Liu, J. Sun, L. Shang, Q. Yang, X. Chen, B. Li, *J. Biophotonics* **2023**, *16*, e202200270.

[15] B. Liu, K. Liu, N. Wang, K. Ta, P. Liang, H. Yin, B. Li, *Talanta* **2022**, *244*, 123383.

[16] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A. A. Bharath, *IEEE Signal Process. Mag.* **2018**, *35*, 53.

[17] S. Y. Hsu, L. R. Yeh, T. B. Chen, W. C. Du, Y. H. Huang, W. H. Twan, M.-C. Lin, Y.-H. Hsu, Y.-C. Wu, H.-Y. Chen, *Molecules* **2020**, *25*, 4792.

[18] H. Li, J. Zhou, Y. Zhou, Q. Chen, Y. She, F. Gao, Y. Xu, J. Chen, X. Gao, *Front. Physiol.* **2021**, *12*, 655556.

[19] H. Sak, A. Senior, F. Beaufays, *arXiv Preprint* arXiv:14021128 **2014**. https://doi.org/10.48550/arXiv.1402.1128

[20] P. A. Gorry, *Anal. Chem.* **2002**, *62*, 570.

[21] Z. M. Zhang, S. Chen, Y. Z. Liang, *Analyst* **2010**, *135*, 1138.

[22] L. Dinh, J. Sohl-Dickstein, S. Bengio, *arXiv Preprint* arXiv: 1605.08803 **2016**. https://doi.org/10.48550/arXiv.1605.08803

[23] J. Burgess, P. O'Kane, S. Sezer, D. Carlin, *Cybersecurity* **2021**, *4*, 15.