

# Conditional Generative Adversarial Network for Spectral Recovery to Accelerate Single-Cell Raman Spectroscopic Analysis

Xiangyun Ma,<sup>||</sup> Kaidi Wang,<sup>||</sup> Keng C. Chou,<sup>\*</sup> Qifeng Li,<sup>\*</sup> and Xiaonan Lu<sup>\*</sup>



Cite This: *Anal. Chem.* 2022, 94, 577–582



Read Online

ACCESS |



Metrics & More

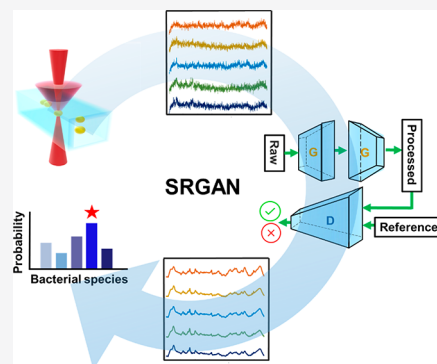


Article Recommendations



Supporting Information

**ABSTRACT:** Raman spectroscopy is a powerful tool to investigate cellular heterogeneity. However, Raman spectra for single-cell analysis are hindered by a low signal-to-noise ratio (SNR). Here, we demonstrate a simple and reliable spectral recovery conditional generative adversarial network (SRGAN). SRGAN reduced the data acquisition time by 1 order of magnitude (i.e., 30 vs 3 s) by improving the SNR by a factor of  $\sim 6$ . We classified five major foodborne bacteria based on single-cell Raman spectra to further evaluate the performance of SRGAN. Spectra processed using SRGAN achieved an identification accuracy of 94.9%, compared to 60.5% using unprocessed Raman spectra. SRGAN can accelerate spectral collection to improve the throughput of Raman spectroscopy and enable real-time monitoring of single living cells.



## INTRODUCTION

Single-cell analysis is an ideal approach to investigate the cell-to-cell heterogeneity in tumor and microbial communities.<sup>1,2</sup> There is an increasing need for analytical methods to comprehensively monitor the phenotypic and physiological changes of single living cells. Raman spectroscopy provides noninvasive and label-free “molecular fingerprints” that have been widely used in biological applications.<sup>3</sup> Different cellular phenotypes have unique molecular compositions that generate subtle variations in their Raman spectra. Raman spectroscopy has the potential to identify compositional changes in individual cells and enable real-time screening at the single-cell level. However, a significant challenge in single-cell analysis using Raman spectroscopy is its low signal-to-noise ratio (SNR). Raman signals are inherently weak ( $\sim 10^{-8}$  scattering probability), especially from a single cell.<sup>4</sup> Thus, the subtle spectral variations from different cells are easily overwhelmed by the background noise, including shot noise, autofluorescence, and detector readout noise.<sup>3</sup> Consequently, spectral recovery is needed to achieve more accurate qualitative or quantitative analysis.

An extended spectral collection time is typically required to achieve a high SNR for Raman spectroscopy. However, a long integration time can cause photodamage living cells and decrease the analytical speed, prohibiting high-throughput single-cell analysis.<sup>5</sup> Thus, reliable computational processing methods are required to rapidly acquire high-quality single-cell Raman spectra. Numerous mathematical techniques have been used to remove the spectral background noise, including polynomial fitting,<sup>6</sup> Savitzky–Golay smoothing,<sup>7</sup> first and second derivatives,<sup>8</sup> wavelet transform,<sup>9</sup> and Fourier trans-

form.<sup>10</sup> However, most of these methods have limitations. For example, polynomial fitting algorithms estimate a baseline using a low-order polynomial that may not be suitable for spectra with high noise. The first and second derivatives require prior knowledge of noise and signal components for peak selection and peak shift, and signal suppression may occur when the SNR is low. Wavelet transform and Fourier transform rely on the frequency difference between signal and noise, but the selection of an appropriate frequency filter is challenging when Raman peaks overlap significantly.

Recently, conditional generative adversarial networks (CGAN) have shown considerable success in various fields,<sup>11</sup> such as image translation, speech enhancement, and video generation. In CGAN, two separate models are trained simultaneously in an adversarial manner, i.e., a *generator* for synthesizing new data and a *discriminator* for distinguishing the generated data from the real ones in the training data set. It provides effective signal enhancement and works end-to-end with no explicit assumptions. However, the potential of CGAN in spectroscopic analysis has not yet been explored.

In this work, we proposed a simple and effective CGAN-based framework named the Spectral Recovery Generative Adversarial Network (SRGAN). The proposed approach directly learns end-to-end mapping between noisy spectra

**Received:** September 30, 2021

**Accepted:** December 28, 2021

**Published:** January 3, 2022



and the corresponding clean spectra in the training data set. Compared to other common denoising methods that focus on removing noise, SRGAN aims to recover the valid Raman signal from the low-SNR spectra. A Raman scattering signal from a single bacterium is inherently weak; therefore, this is an excellent target to demonstrate the effectiveness of our spectral recovery method. This approach was validated to classify leading foodborne bacteria, including *Salmonella*, *Escherichia*, *Listeria*, *Staphylococcus*, and *Campylobacter*. It can be combined with Raman spectroscopy to achieve high-throughput and real-time monitoring of living bacterial cells. To the best of our knowledge, this is the first approach of using CGAN for Raman spectral processing.

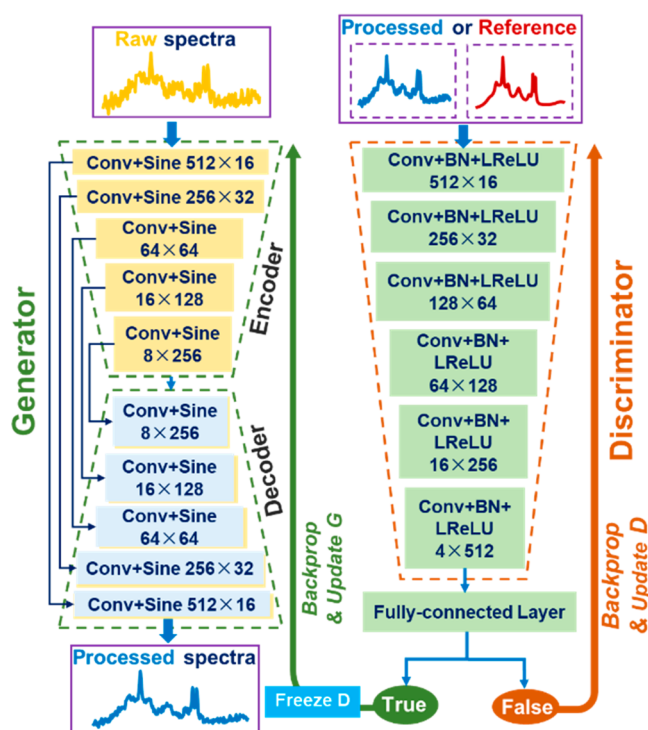
## EXPERIMENTAL SECTION

**Sample Preparation.** Single-cell Raman spectra from bacteria were used to validate the proposed SRGAN. Five leading foodborne bacteria were investigated, including *Campylobacter jejuni* F38011, *Escherichia coli* K12, *Listeria monocytogenes* ATCC19113, *Staphylococcus aureus* Newman, and *Salmonella enterica* serovar Typhimurium SL1344. An overnight bacterial culture was individually harvested at an early stationary phase, and the bacterial cells were resuspended in phosphate buffered saline (PBS) for spectral collection. The detailed information on bacterial cultivation and sample preparation for spectral collection is available in the [Supporting Information \(SI\)](#).

**Raman Spectral Collection.** Single-cell Raman spectra were collected using a home-built confocal Raman spectroscopic system with a 671 nm laser. Bacteria were introduced using a polydimethylsiloxane (PDMS)-based microfluidic chip. All the parameters of system configuration and spectral collection are described in the [SI](#). Single-cell Raman spectra with low SNR were collected with an integration time of 3 s. Another set of Raman spectra of the same cells were integrated for 30 s and used as the reference spectra with a high SNR. For each bacteria, ~200 individual single cells were measured from three independent batches to minimize the influence of biological variability, environmental noise, and experimental operations.

**Model Construction.** We introduced SRGAN as a general-purpose solution to improve the SNR of single-cell Raman spectra.<sup>11,12</sup> The SRGAN included a generator (G) and a discriminator (D) ([Figure 1](#)). Raw spectra collected with an integration time of 3 s were used as the input data. The objective of G was to generate processed Raman spectra with an improved SNR from the input data so that it became challenging for D to distinguish the generated spectra from the corresponding real spectra (30 s integration time). Then, D and G were trained jointly and gradually with the training data sets. Briefly, G combined an encoder and a decoder structure with skip layers. The encoder was composed of five blocks that contained a one-dimensional convolutional layer (Conv) with a filter size of  $5 \times 1$  and a sinusoidal activation function (Sine). The stride was set to 2 with padding, and the number of extracted feature maps in each layer is shown in [Figure 1](#). The decoder of G had the same number of blocks as the encoder. It reversed the down-sampling process and output a processed spectrum of the same size as that of the input spectrum. The skipping connections could directly send the information (e.g., the spectral energy distribution) to the decoder.

The discriminator D was trained to differentiate the processed spectrum (generated by G) and the input reference



**Figure 1.** Architecture of the proposed conditional generative adversarial network for spectral recovery (SRGAN). A generator (G) and a discriminator (D) are trained jointly and gradually in SRGAN. G serves as a mapping function to recover a processed Raman spectrum from an input raw spectrum, while D is trained to differentiate the processed spectrum (generated by G) from the corresponding reference spectrum. G includes an encoder and a decoder structure, and each has five blocks. Each block contains a one-dimensional convolutional layer (Conv), a sinusoidal activation function (Sine), and different numbers of extracted feature maps. D consists of six blocks with a Conv-batch normalization layer and a leaky rectified linear unit (Conv + BN + LReLU), followed by a fully connected layer.

spectrum. D started with five Conv-batch normalization layers and a leaky rectified linear unit (Conv + BN + LeakyReLU) blocks, followed by a fully connected layer. The loss functions in the SRGAN were defined as

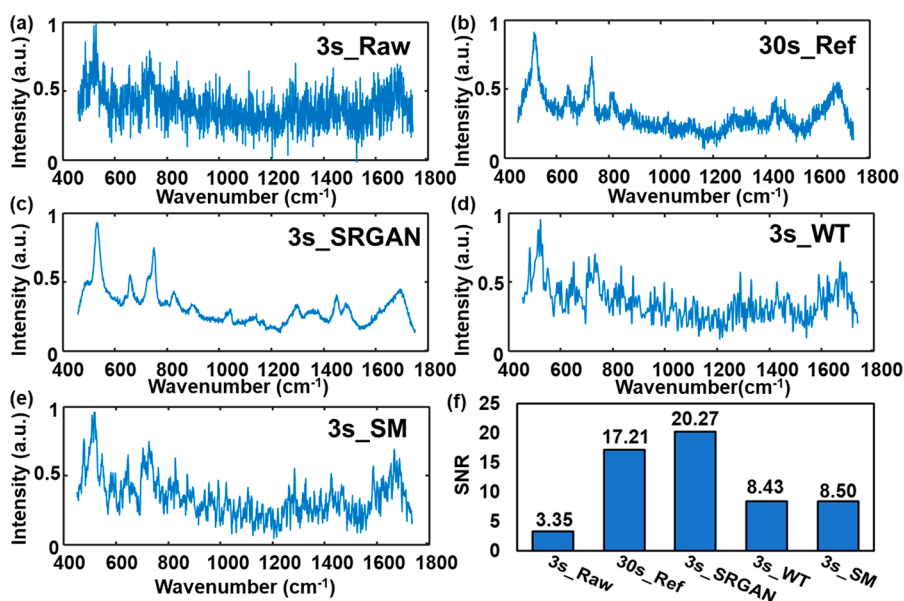
$$L_G = \frac{1}{N} \sum_{n=1}^N (D(G(s^n), s^n) - 1)^2 + \lambda \left\| G(s^n) - x^n \right\|_1 \quad (1)$$

$$L_D = \frac{1}{N} \sum_{n=1}^N (D(x^n, s^n) - 1)^2 + (D(G(s^n), s^n))^2 \quad (2)$$

where  $s^n$  is the raw spectrum,  $x^n$  is the reference spectrum with  $n \in N$ ,  $N$  is the number of training spectra, and  $\lambda$  is the hyperparameter for the magnitude of the L1 norm. We selected the L1 norm as it could improve the SNR of the processed spectra.<sup>13</sup> The least-squares function with binary coding was introduced to resolve the vanishing gradient problem.<sup>14</sup> The model was trained using RMSprop for 100 epochs with a learning rate of 0.0001 and an effective batch size of 50. Python was used to run SRGAN.

A large and comprehensive training data set increases the robustness and stability of the model. In this study, the training data set consists of three parts:

1. **Spectral Database for Organic Compounds.** A total of 600 Raman spectra of organic compounds were randomly



**Figure 2.** Average raw, reference, and processed single-cell Raman spectra of *Campylobacter jejuni*. (a) Raw Raman spectra of *C. jejuni* with an integration time of 3 s. (b) Reference Raman spectra with an integration time of 30 s. (c) Processed Raman spectra using the proposed spectral recovery generative adversarial network (SRGAN). (d) Processed Raman spectra using wavelet transformation (WT). (e) Processed Raman spectra using smoothing (SM). (f) Signal-to-noise ratio (SNR) of all the raw, reference, and processed Raman spectra.

selected from the Spectral Database for Organic Compounds<sup>15</sup> to represent abundant organic compositions of bacterial cells. We artificially added different types of noise to those spectra following the procedures mentioned in a previous study<sup>16</sup> to mimic the background noises found in the real-world spectra. These spectra with noises serve as the input data set for G, and the corresponding clean spectra are used as the reference data set for D.

**2. Real-World Data Set.** Single-cell Raman spectra were collected from bacteria using our home-built Raman spectrometer as aforementioned, including all types of noises in our experimental settings.

**3. Simulated Data Set.** Due to the difficulty in obtaining a high number of standard spectra, we also constructed a large synthetic data set to extend the diversity of the training data set. The simulated spectra consist of Lorentz and Gaussian peaks with random intensity, line width, peak number, and distribution. The simulated spectra were also artificially degraded with multiple types of noises.

**Classification of Different Bacteria.** To assess the performance of SRGAN, classification of different foodborne bacteria based on single-cell Raman spectra was conducted. Raman spectra from different bacteria were collected with the same sample preparation and data acquisition procedures to reduce experimental variations. A machine learning technique, namely, convolutional neural network (CNN), was used to discriminate different bacteria. The construction (Figure S1) and validation of CNN is detailed in the SI.

## RESULTS AND DISCUSSION

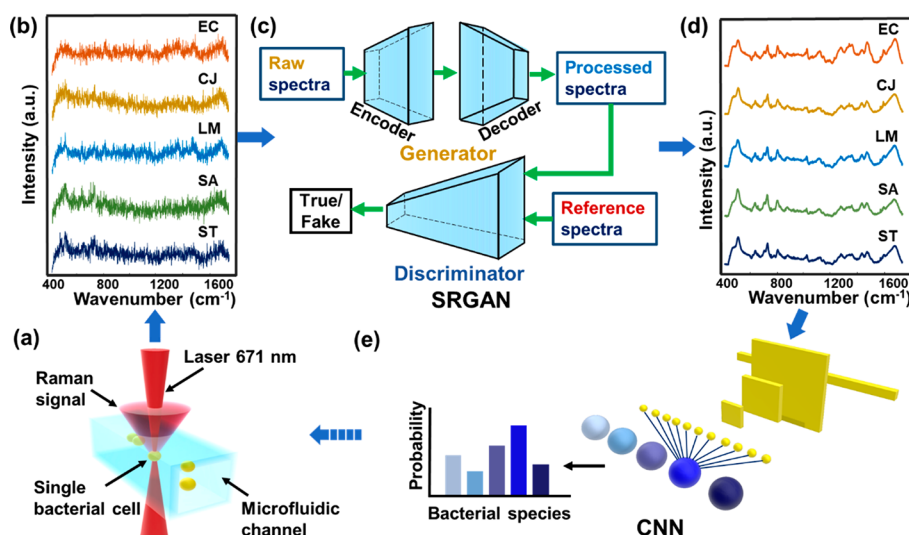
Raman spectroscopy for single cell analysis has been widely used in biology as it is rapid, label free, and nondestructive. Noise removal is considered as a primary and inevitable step for Raman spectroscopic data analysis, especially for single-cell Raman spectra with low SNR. Elimination of various noises in the experimentally obtained Raman spectra can improve the spectral quality and thus increase the accuracy of the

subsequent data analyses. As SNR is positively correlated to integration time, improvement of SNR can lead to the reduction of spectral acquisition time. In this study, we presented a deep learning technique SRGAN to transform noisy raw Raman spectra into high-quality spectra with improved SNRs. This approach can significantly accelerate single-cell Raman spectroscopic analysis and has the potential to expand its application in real-time analysis.

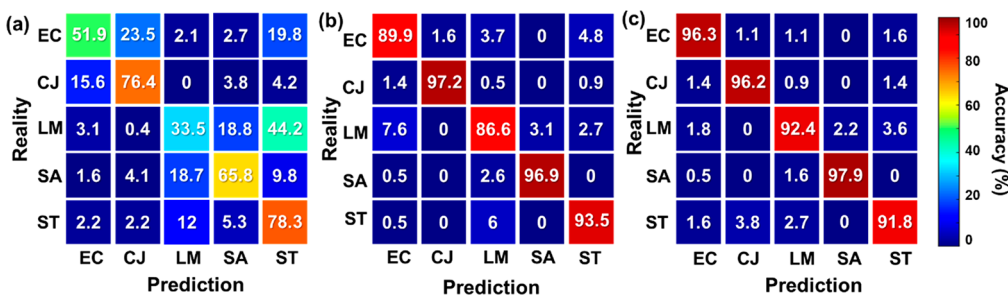
We evaluated the performance of SRGAN using single-cell Raman spectra from *Campylobacter*, a leading bacterial cause of gastrointestinal diseases.<sup>17</sup> Raw Raman spectra collected with an integration time of 3 s were used as input. Raman spectra acquired with an integration time of 30 s were employed as reference spectra as they have better spectral quality and have been commonly used in other single-cell Raman spectroscopic studies.<sup>18,19</sup> As indicated in Figure 2a, a raw single-cell Raman spectrum shows inferior spectral quality. Specifically, Raman peaks are overwhelmed by noise, and no clear spectral features can be observed. In comparison, unique Raman spectral patterns can be identified in the spectra processed using SRGAN (Figure 2c). These peaks are attributed to biochemical compositions of bacterial cells and the PDMS chip (Figure S2). Compared to reference Raman spectra (Figure 2b), the processed Raman spectra recovered abundant peak information, and the spectral patterns are mostly preserved without peak shift.

SNR is an important parameter to quantitatively evaluate the quality of Raman spectra.<sup>20</sup> In this study, SNR is defined as the average peak intensity divided by the standard deviation of the background. A long integration time is usually required to improve the SNR of Raman spectra. The SNR of raw Raman spectra increased from 3.35 to 17.21 with increasing integration times of 3, 5, 10, and 30 s (Figure S3). After spectral recovery using SRGAN, the SNR of the 3 s measurement was improved to 20.27, which was higher than that of the reference Raman spectra. Therefore, data acquisition time can be reduced at least 1 order of magnitude





**Figure 3.** Scheme of experimental design. (a) Single-cell Raman spectra were collected using a homemade confocal Raman spectroscopic system combined with a microfluidic chip. (b) Raw Raman spectra with an integration time of 3 s for five leading foodborne bacteria. (c) Framework of SRGAN. (d) Processed Raman spectra using SRGAN. (e) Bacterial classification using a convolutional neural network (CNN). CJ, *Campylobacter jejuni* F38011; EC, *Escherichia coli* K12; LM, *Listeria monocytogenes* ATCC19113; SA, *Staphylococcus aureus* Newman; ST, *Salmonella enterica* serovar Typhimurium SL1344.



**Figure 4.** Classification results of five leading foodborne bacteria using a convolutional neural network (CNN). Confusion matrices for raw Raman spectra with an integration time of 3 s (a), processed spectra recovered using SRGAN (b), and reference Raman spectra with an integration time of 30 s (c). The overall accuracies are 60.5%, 92.7%, and 94.9%, respectively. CJ, *Campylobacter jejuni* F38011; EC, *Escherichia coli* K12; LM, *Listeria monocytogenes* ATCC19113; SA, *Staphylococcus aureus* Newman; ST, *Salmonella enterica* serovar Typhimurium SL1344.

(i.e., from 30 to 3 s), and enhanced spectral quality is achieved using SRGAN.

A comparative study was conducted between SRGAN and two other frequently used denoising methods, namely, wavelet transformation and smoothing. The quality of processed Raman spectra (Figure 2d and e) was improved compared to the raw Raman spectra, but a considerable number of spectral features were still overwhelmed by the noise. The SNRs of the processed Raman spectra using wavelet transformation and smoothing were 8.50 and 8.43, respectively, which were lower than 17.21 of SRGAN (Figure 2f). Wavelet transformation and smoothing have been widely used for noise removal in a wide range of applications. However, they have limitations in certain situations. The challenges in selecting appropriate wavelet filters, wavelet threshold, and other parameters in wavelet transform may affect its denoising results.<sup>9</sup> For smoothing, there is a trade-off between spectral resolution and smoothing outcome as the weak spectral features might be distorted when the window size is large.<sup>7</sup> The superior performance of SRGAN might be attributed to its ability to learn and imitate the distribution of reference data and synthetically generate examples that optimally resemble the real data. Also, multiple types of noise included in the

training data sets contribute to the robustness of SRGAN for different applications. Taken together, SRGAN showed a better performance in recovering single-cell Raman spectra with a low SNR than wavelet transformation and smoothing.

We further performed the classification of various bacteria to test the performance of SRGAN. A total of 1000 single-cell Raman spectra from five leading foodborne bacteria, including *C. jejuni*, *E. coli*, *L. monocytogenes*, *S. aureus*, and *S. typhimurium*, were collected using our home-built Raman spectroscopic system with a microfluidic device (Figure 3a). These bacteria are responsible for a large number of food-associated events, such as foodborne illnesses.<sup>21</sup> Rapid detection of bacteria at an early stage is crucial to ensure biomedical diagnostics and food safety. Raman spectroscopy has been used in the detection and identification of foodborne bacteria as it is rapid, label free, and easy to perform. However, it is technically challenging to distinguish different bacterial populations based on single-cell Raman spectra because the weak signals from single cells can be easily suppressed by background noise. Average raw Raman spectra recorded from five bacterial genera with an integration time of 3 s are displayed in Figure 3b. After being recovered by SRGAN (Figure 3c), the processed spectra indicate improved spectral quality (Figure 3d). A CNN model (Figure 3e) was

applied to classify bacteria using the raw, processed, and reference Raman spectra (30 s integration), respectively, and the identification accuracies are summarized in Figure 4 and Table S1. The overall accuracy for the discrimination of five bacteria was 60.5% using raw spectra. As Raman signals from bacteria are largely masked by background noise in the raw Raman spectra, minor variations among different bacterial cells cannot be distinguished, leading to low identification accuracy. The identification accuracy of the recovered spectra could be improved to 94.9% which was higher than 92.7% of the reference Raman spectra. By using SRGAN, data acquisition time was reduced over 10 times (i.e., 3 vs 30 s) compared to of the typical Raman spectroscopic studies for bacterial identification, and the identification accuracy was enhanced as well. The superior accuracy of the processed Raman spectra further confirms that SRGAN can preserve most original spectral features from bacterial compositions without introducing artificial interference during the recovery process.

CGAN is a relatively new deep learning technique that has been increasingly recognized in speech enhancement, image generation, and style transfer.<sup>22</sup> The main advantages of the proposed SRGAN are as follows: (1) The optimization process is rapid. Without an iterative optimization procedure, the entire enhancement process usually takes less than a few milliseconds on a modern computer. (2) No parameter adjustment is required. All the parameters have been fixed in the training process, which can deal with various noise types in the training data set. (3) It has a superior recovery capability for Raman spectra with a low SNR. SRGAN is more robust to extracting weak signals compared to the traditional noise removal methods. On the basis of these advantages, SRGAN can be integrated with single-cell sorting techniques, such as a Raman optical tweezer, to allow simultaneous spectral analysis and cell manipulation. This system has the potential to achieve accurate and real-time monitoring of single living cells with high throughput and opens a new window for drug screening, bacterial detection, and cancer diagnosis.

## CONCLUSIONS

We showed a novel and reliable approach to recover single-cell Raman spectra with a low SNR. Comprehensive training data sets were included to increase model robustness and reduce artifacts. SNR of Raman spectra of single bacterial cells can be increased over 6 times, and the processed Raman spectra recovered most spectral features compared to the reference spectra. Our approach demonstrated a superior performance in comparison to other commonly used denoising methods. For the classification of five major foodborne bacteria, the SRGAN-recovered Raman spectra achieved an improved identification accuracy compared to the unprocessed Raman spectra (i.e., 94.9% vs 60.5%). SRGAN accelerates data acquisition time over 10 times (i.e., 30 vs 3 s), significantly improving speed and throughput of Raman spectroscopy for single-cell analysis. It can also be generalized to other spectroscopic techniques, including nuclear magnetic resonance, mass spectrometry, and infrared spectroscopy to improve spectral quality and analysis speed.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.1c04263>.

Additional experimental details, structure of convolutional neural network, Raman spectrum of PDMS chip, raw Raman spectra of *C. jejuni* with different integration times, and classification results of five bacteria (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**Keng C. Chou** – Department of Chemistry, The University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada; [orcid.org/0000-0002-8782-5253](https://orcid.org/0000-0002-8782-5253); Email: [kcchou@chem.ubc.ca](mailto:kcchou@chem.ubc.ca)

**Qifeng Li** – School of Precision Instrument and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China; [orcid.org/0000-0001-8813-5054](https://orcid.org/0000-0001-8813-5054); Email: [qfli@tju.edu.cn](mailto:qfli@tju.edu.cn)

**Xiaonan Lu** – Department of Food Science and Agricultural Chemistry, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, Quebec H9X 3V9, Canada; [orcid.org/0000-0003-0254-0345](https://orcid.org/0000-0003-0254-0345); Email: [xiaonan.lu@mcgill.ca](mailto:xiaonan.lu@mcgill.ca)

### Authors

**Xiangyun Ma** – School of Precision Instrument and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China; Department of Chemistry, The University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

**Kaidi Wang** – Department of Food Science and Agricultural Chemistry, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, Quebec H9X 3V9, Canada

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.1c04263>

### Author Contributions

<sup>†</sup>X. Ma and K. Wang contributed equally. X. Ma, K. Wang, and X. Lu designed this study. K. Wang collected the data. K. C. Chou designed and built the Raman microscope. K. Wang and X. Ma analyzed and described the data. Q. Li provided critical comments on model construction. K. Wang, X. Ma, and X. Lu contributed to the writing.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada in the form of a Discovery Grant (NSERC RGPIN-2019-03960) and a Discovery Accelerator Grant (NSERC RGPIN-2019-00024) to Xiaonan Lu and an On-going New Opportunities Fund (CFI project 10499) and a Discovery Grant (NSERC RGPIN-2019-05509) to Keng C. Chou, as well as a National Natural Science Foundation of China (NSFC22174098) and a Key Research and Development Program of Tianjin (20YFZCSN00530) to Qifeng Li.

## REFERENCES

- (1) He, Y.; Wang, X.; Ma, B.; Xu, J. *Biotechnology Advances* **2019**, *37*, 107388.
- (2) Zhou, Y.; Bian, S.; Zhou, X.; Cui, Y.; Wang, W.; Wen, L.; Guo, L.; Fu, W.; Tang, F. *Cancer Cell* **2020**, *38*, 818–828.e5.
- (3) Butler, H. J.; Ashton, L.; Bird, B.; Cinque, G.; Curtis, K.; Dorney, J.; Esmonde-White, K.; Fullwood, N. J.; Gardner, B.; Martin-Hirsch, P.

- L.; Walsh, M. J.; McAinsh, M. R.; Stone, N.; Martin, F. L. *Nat. Protoc.* **2016**, *11*, 664.
- (4) Ellis, D. I.; Goodacre, R. *Analyst* **2006**, *131*, 875–885.
- (5) Ho, C.-S.; Jean, N.; Hogan, C. A.; Blackmon, L.; Jeffrey, S. S.; Holodniy, M.; Banaei, N.; Saleh, A. A.; Ermon, S.; Dionne, J. *Nat. Commun.* **2019**, *10*, 1–8.
- (6) Wang, Z.; Zhang, M.; Harrington, P. d. B. *Anal. Chem.* **2014**, *86*, 9050–9057.
- (7) Vivó-Truyols, G.; Schoenmakers, P. J. *Anal. Chem.* **2006**, *78*, 4598–4608.
- (8) Xie, Y.; Yang, L.; Sun, X.; Wu, D.; Chen, Q.; Zeng, Y.; Liu, G. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2016**, *161*, 58–63.
- (9) Ehrentreich, F.; Sümmchen, L. *Anal. Chem.* **2001**, *73*, 4364–4373.
- (10) Koo, I.; Zhang, X.; Kim, S. *Anal. Chem.* **2011**, *83*, 5631–5638.
- (11) Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017; pp 1125–1134.
- (12) Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 2014.
- (13) Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A. A. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016; pp 2536–2544.
- (14) Mao, X.; Li, Q.; Xie, H.; Lau, R. Y.; Wang, Z.; Paul Smolley, S. Least squares generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017; pp 2794–2802.
- (15) Tanabe, K.; Hiraishi, J. *Spectral Database for Organic Compounds*, 2018. <https://sdbs.db.aist.go.jp> (accessed December 2021).
- (16) Smulko, J. M.; Dingari, N. C.; Soares, J. S.; Barman, I. *Bioanalysis* **2014**, *6*, 411–421.
- (17) Burnham, P. M.; Hendrixson, D. R. *Nature Reviews Microbiology* **2018**, *16*, 551–565.
- (18) Fang, T.; Shang, W.; Liu, C.; Liu, Y.; Ye, A. *Anal. Chem.* **2020**, *92*, 10433–10441.
- (19) Li, M.; Xu, J.; Romero-Gonzalez, M.; Banwart, S. A.; Huang, W. E. *Curr. Opin. Biotechnol.* **2012**, *23*, 56–63.
- (20) Li, Q.; Ma, X.; Wang, H.; Wang, Y.; Zheng, X.; Chen, D. *Opt. Express* **2018**, *26*, 525–530.
- (21) Yan, S.; Wang, S.; Qiu, J.; Li, M.; Li, D.; Xu, D.; Li, D.; Liu, Q. *Talanta* **2021**, *226*, 122195.
- (22) Kench, S.; Cooper, S. J. *Nature Machine Intelligence* **2021**, *3*, 299–305.