



Article

A Deep-Learning Based System for Rapid Genus Identification of Pathogens under Hyperspectral Microscopic Images

Chenglong Tao ^{1,2,3} , Jian Du ^{1,3}, Yingxin Tang ⁴ , Junjie Wang ^{1,2}, Ke Dong ⁵, Ming Yang ⁵, Bingliang Hu ^{1,3,*} and Zhoufeng Zhang ^{1,3,*}

¹ Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China; chengtao@126.com (C.T.); dujian@opt.ac.cn (J.D.); gjswj@gmail.com (J.W.)

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Key Laboratory of Biomedical Spectroscopy of Xi'an, Xi'an 710119, China

⁴ Independent Researcher, Changsha 410000, China; yxtang1995@163.com

⁵ The Second Affiliated Hospital of Air Force Military Medical University, Xi'an 710119, China; drdongke@163.com (K.D.); yangming745@163.com (M.Y.)

* Correspondence: hbl@opt.ac.cn (B.H.); zhangzhoufeng@opt.ac.cn (Z.Z.)

Abstract: Infectious diseases have always been a major threat to the survival of humanity. Additionally, they bring an enormous economic burden to society. The conventional methods for bacteria identification are expensive, time-consuming and laborious. Therefore, it is of great importance to automatically rapidly identify pathogenic bacteria in a short time. Here, we constructed an AI-assisted system for automating rapid bacteria genus identification, combining the hyperspectral microscopic technology and a deep-learning-based algorithm Buffer Net. After being trained and validated in the self-built dataset, which consists of 11 genera with over 130,000 hyperspectral images, the accuracy of the algorithm could achieve 94.9%, which outperformed 1D-CNN, 2D-CNN and 3D-ResNet. The AI-assisted system we developed has great potential in assisting clinicians in identifying pathogenic bacteria at the single-cell level with high accuracy in a cheap, rapid and automatic way. Since the AI-assisted system can identify the pathogenic genus rapidly (about 30 s per hyperspectral microscopic image) at the single-cell level, it can shorten the time or even eliminate the demand for cultivating. Additionally, the system is user-friendly for novices.

Keywords: infectious pathogens; hyperspectral microscopy; bacteria identification; artificial intelligence; imaging; genus; spectral characteristics



Citation: Tao, C.; Du, J.; Tang, Y.; Wang, J.; Dong, K.; Yang, M.; Hu, B.; Zhang, Z. A Deep-Learning Based System for Rapid Genus Identification of Pathogens under Hyperspectral Microscopic Images. *Cells* **2022**, *11*, 2237. <https://doi.org/10.3390/cells11142237>

Academic Editors: An Pan, Baoli Yao, Chao Zuo, Fei Liu, Jiamiao Yang, Liangcai Cao and Yu Xue

Received: 12 May 2022

Accepted: 15 July 2022

Published: 19 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Infectious diseases are a significant threat to public health and remain the leading cause of death in developing countries, despite remarkable medical progress. Therefore, how to prevent and treat them efficiently is of great significance [1–4].

Currently, there are several conventional methods for bacteria detection, such as conventional morphological examination [5], polymerase chain reaction (PCR) [6], the matrix assisted Laser Desorption Ionization Time of Flight (MALDI-TOF) [7], biochemical identification [8], serological tests [9], etc. However, a morphological examination is time-consuming, since pathogens must be cultivated to be colonies. PCR, which needs skillful operations, is only adopted for emergencies due to the high expense. Especially, MALDI-TOF is regarded as the “gold standard”. However, it is scarcely used in clinics because of the complicated operation and expensive cost. Additionally, the reproducibility of results is poor due to the destruction to the specimen. As for patients, broad-spectrum antibiotics are first-line-adopted when the cause of infection is unclear. Unfortunately, the treatment seldom brings a satisfactory efficacy. Moreover, it might also bring a heavy financial and psychological burden to the patient and potential risk of worsening the condition [10,11].

Due to the relative scarcity of laboratory physicians, there is an urgent need to rapidly automate pathogenic bacteria identification at a low price.

By utilizing spectral profiles of bacterial suspensions, Dimitris et al. identified five bacterial species [12] and demonstrated that spectral technique could be a rapid, cheap and noninvasive tool for bacterial identification. However, their method has only spectral information and limited recognition capability. To adapt to more complex recognition scenarios, hyperspectral technology (finer spectral resolution than a spectral technique) and imaging technology are combined into hyperspectral imaging (HSI) technology, which was used to identify bacteria with colony hyperspectral images [13–16]. Since hyperspectral images have three dimensions, similar to a cube, they are also called datacubes.

However, these methods are colony-based and require the culturing of bacteria to a considerable amount, which is very time-consuming. For faster identification of pathogens, HSI and microscopic techniques were combined to develop microscopic hyperspectral imaging (HMI). HMI was exploited to detect food-borne pathogens at the cell level [17–20]. However, the narrow range of wavelengths (450–800 nm) and coarse spectral resolution (4 nm) are insufficient for classifying infectious pathogens. In addition, a microscopic datacube was decomposed or averaged, which destroyed the strong correlations between the spectral and spatial information of the datacube. Although convolutional neural networks (CNNs) were utilized, the small-scale dataset was inadequate for an excellent CNN-based algorithm and clinical practice.

In this study, a new HMI system with a wider wavelength range (440–1023 nm) and finer spectral resolution (2.1 nm) was developed. Over 130,000 HMIs (hyperspectral microscopic images) for bacteria of 11 genera were collected with this equipment. Moreover, a 3D-CNN-based network was proposed, called the Buffer Net, to maintain the data integrity in HMIs. For noninvasive, fast and accurate bacteria detection, the HMI system and trained Buffer Net were integrated to be an AI-assisted system.

In Section 2, we detail the collection and processing of the data, system design and integration. The results are shown in Section 3, and the significance of our work is discussed in Section 4. In Section 5, we concisely conclude the paper.

2. Materials and Methods

The workflow diagram of the AI-assisted system is shown in Figure 1. Additionally, visualization of the cropping individual bacteria and the main functions of the system are included.

2.1. Bacteria Strains

Most strains in this study were from The Second Affiliated Hospital of Air Force Military Medical University. In addition, in order to increase the diversity of our data, bacteria from the American Type Culture Collection (ATCC) were also adopted in the study. The genera used in this study are summarized in Table 1. The extracted infectious bacteria were purified (if necessary) and incubated on blood plates at a constant temperature of 35–37 °C for 24–48 h. Afterwards, some of the individual colonies were dissolved in turbidity tubes under high pressure with saline (that is, to sterilize). A slide was taken out to have the bacterial suspension in the inoculation loop coated on it evenly. Then, the microscopic slide was dried naturally in the biosafety cabinet. Finally, a 75% alcohol solution was applied to the dried slide, and then, absorbent paper was used to dry the slide. After implementing the above process, the slide was well-prepared for observation and imaging under the HMI system.

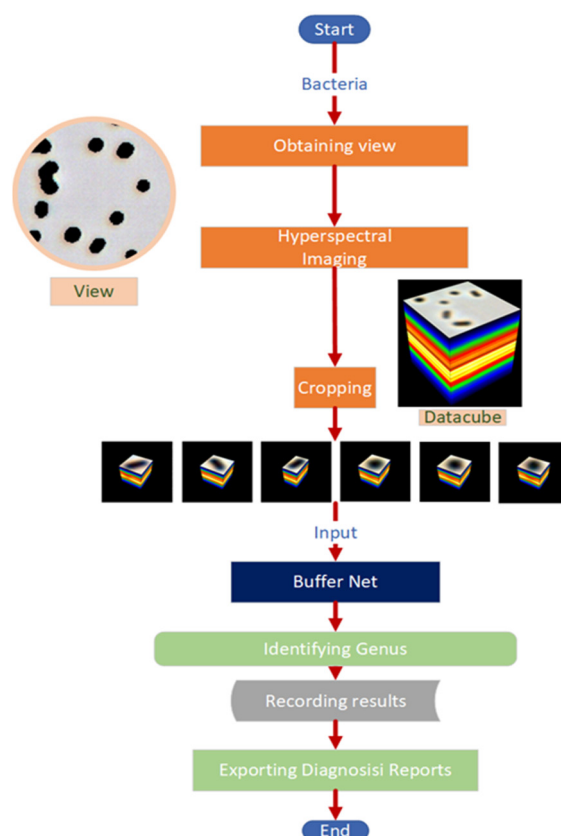


Figure 1. Workflow diagram of the AI-assisted system. The field of view of the bacteria made into slides is acquired under a microscope with 100× magnification and then imaged. The bacteria in the datacube obtained by imaging are cropped out and fed into the Buffer Net. The identification results are recorded and exported to help physicians make a diagnosis.

Table 1. The number of bacteria.

Genera	Numbers for Training	Numbers for Testing	Total
Stenotrophomonas	9070	3888	12,958
Escherichia	7135	3058	10,193
Morganella	7384	3165	10,549
Burkholderia	6337	2717	9054
Serratia	8126	3483	11,609
Pseudomonas	8954	3838	12,792
Acinetobacter	7185	3080	10,265
Klebsiella	11,255	4825	16,080
Proteus	11,435	4902	16,337
Staphylococcus	8182	3508	11,690
Enterococcus	8887	3810	12,697
total	93,950	40,274	134,224

2.2. Hyperspectral Microscopic Imaging (HMI) System

A built-in push-broom HMI system (Figure 2a) was developed, which included a microscope and hyperspectral imaging system. An Olympus Biomedical BX43 upright microscope (Olympus, Tokyo, Japan) was used in the system. A 30-W halogen lamp was utilized as the light source. The objective parameters for bright field imaging: 100×/0.90. The model number of the sensor was acA2000-165 umNIR (Basler, Ahrensburg, Germany). The spatial resolution of our hyperspectral camera was about 0.67 μm, the average spectral interval was 2.1 nm, and the field of view (FOV) was 64 × 64 μm². Then, a datacube was obtained containing 277 bands (440–1023 nm).

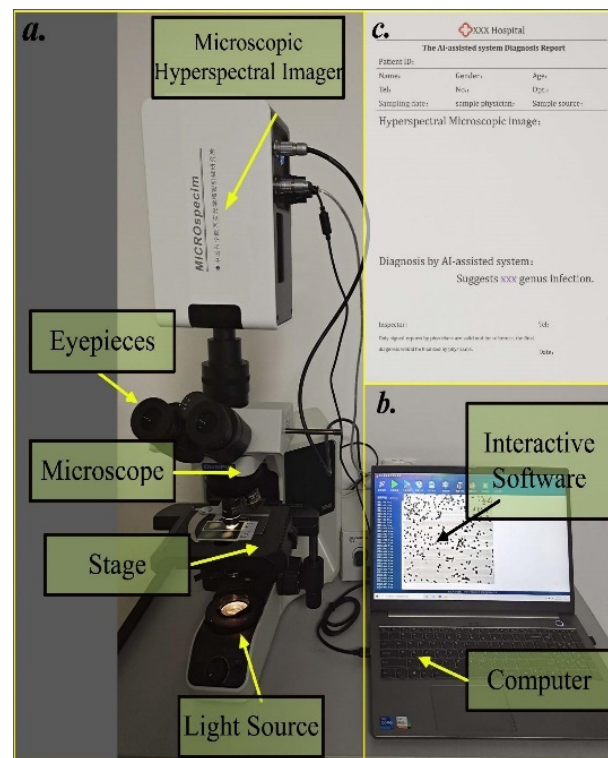


Figure 2. The components of the integrated AI-assisted system. (a) The HMI system. It is composed of a microscope and hyperspectral camera and is responsible for detecting and imaging the view containing bacteria. (b) The host, where data storage processing, bacteria identification and interaction are deployed. (c) A template for the PDF files, i.e., the diagnostic report, output by the AI-assisted system.

2.3. Data Collection and Preprocessing

Due to the low signal-to-noise ratios of the front and rear bands, the first 50 and the last 27 channels of the 277 bands were discarded. To even out the brightness of the entire image, each scanning column was averaged, and then, the means were divided by each pixel in the scanning column. To avoid the influence of extreme values, the 300 maximum and 300 minimum pixels in the scanning column were excluded when calculating the average. This process is called flat-field correction [21].

Then, a clustering algorithm, K-means [22], was adopted to separate the bacterium from the background in HMIs, and a binary image was obtained. The background was calculated by pixel-wise multiplication on the binary image and the whole datacube. Then, the average background brightness was then calculated.

The image where pixel values were like transmittance was obtained by dividing the average background brightness band-wise. The standardization is in order to eliminate the effects of different lighting. The bacterium was accurately identified by its location and size in the image by finding the connected area in the binary image. Then, the bacteria were selected and cropped from the $200 \times 1000 \times 1000$ datacube (Figure 1). The connected areas at the boundary were discarded due to not being intact. The spectral dimensions of the images obtained for individual bacterium was 200, and there was no fixed value for the length and width of the images due to the different pathogen sizes. Finally, more than 130,000 HMIs of 11 infectious genera were collected and randomly assigned to the training set and test set at 70% and 30%, respectively. The number of samples is presented in Table 1.

2.4. Model Design

2.4.1. Deep Learning

Deep learning aims to design a deep network f and determine its parameters θ and to fit the relationship between the input data x and output y^{pre} as accurately as possible. Deep networks generally consist of three parts: input layer, hidden layers and output layer, all of which are composed of neurons. The input layer is only responsible for the input of data, and the data is propagated in the hidden layer with linear and nonlinear transformations. This is also the process of extracting data features. Finally, the data x is mapped as y^{pre} and output from the output layer. This whole process is the forward propagation of the network. y^{pre} and x correspond to the ground truth y^t to calculate the loss. The loss is input into f from the output layer through the hidden layer. This process is the backward propagation of the neural network. During training, after iterating forward and backward propagation, the prediction y^{pre} for x would get closer and closer to the true value y^t . When the condition is satisfied, the training is stopped, and the parameter θ is fixed. The trained neural network f can be used to predict the y^{test} value corresponding to the input data x^{test} .

2.4.2. Buffer Net

As mentioned above, HMIs have three dimensions: spectrum, height and width. The spatial information included in the height and width dimensions is strongly related to the spatial information in the spectral dimension. Extracting the features of HMIs without losing their correlation is extremely important.

Since the 1D-CNN filter is only able to extract the features of one-dimensional spectral profiles and 2D-CNN extracts the spatial features of two-dimensional images, it is not possible for them to extract spectral-spatial features in HMIs [23–25]. Therefore, 3D-CNN uses three-dimensional (3D) filters that are suitable to extract the features of datacubes [26]. For a datacube, $D \in R^{S \times H \times W}$, 3D-CNN uses a 3D filter $F \in R^{K \times I \times J}$ to execute the convolution operation. A 3D feature map $M \in R^{U \times V \times Z}$ is obtained. Assuming the stride is t , $M[s, h, w]$ can be computed as Equation (1) when it is a value located at (s, h, w) in M . Traverse the image with F , and finally, output M . Since D is not decomposed or averaged, the correlation between the spectral and spatial information is retained.

$$\begin{aligned} M[s, h, w] &= (D * F)[s \times t, h \times t, w \times t] \\ &= \sum_{k=1}^K \sum_{i=1}^I \sum_{j=1}^J D[s \times t - k, h \times t - i, w \times t - j] \times F[k, i, j] \end{aligned} \quad (1)$$

Bacteria in datacubes are small objects with resolutions less than 32×32 pixels (defined in [27]), and the resolution of their feature maps will be less than 1 pixel in the feature map after several down samplings. Therefore, two kinds of convolutional layers were designed: a down sampling layer and buffering layer. For expanding the receptive field of the neurons, the stride of the down sampling was set as 2 to down sample the datacube. The stride of the buffering layers is fixed to be 1 for enhancing the representative abilities without reducing the resolution of the feature maps. The convolutional kernels of both layers were $3 \times 3 \times 3$ for extracting detailed information of the datacubes.

A block was composed of a down sampling layer and three buffering layers. In the Buffer Net (Figure 3), a total of four blocks were included. Moreover, the spectral dimension was compressed when the HMI images came into the network by setting the first filter to down sample images on the spectral dimension with a stride of (2, 1, 1) and a convolutional kernel (5, 1, 1).

The dimensions of the feature maps output by the final block were $7 \times 2 \times 2$. The feature maps were flattened into a vector and input to the SoftMax function through 1536 full connections after a 3D average pooling, which would transfer the 11-way output into eleven probabilities for each bacterial genus, and the final output was the genus with the highest probability. Our network consisted of 17 3D convolutional layers (1 + 4 down sampling layers, 4 × 3 buffering layers), 1 fully connected layer and 1 SoftMax layer.

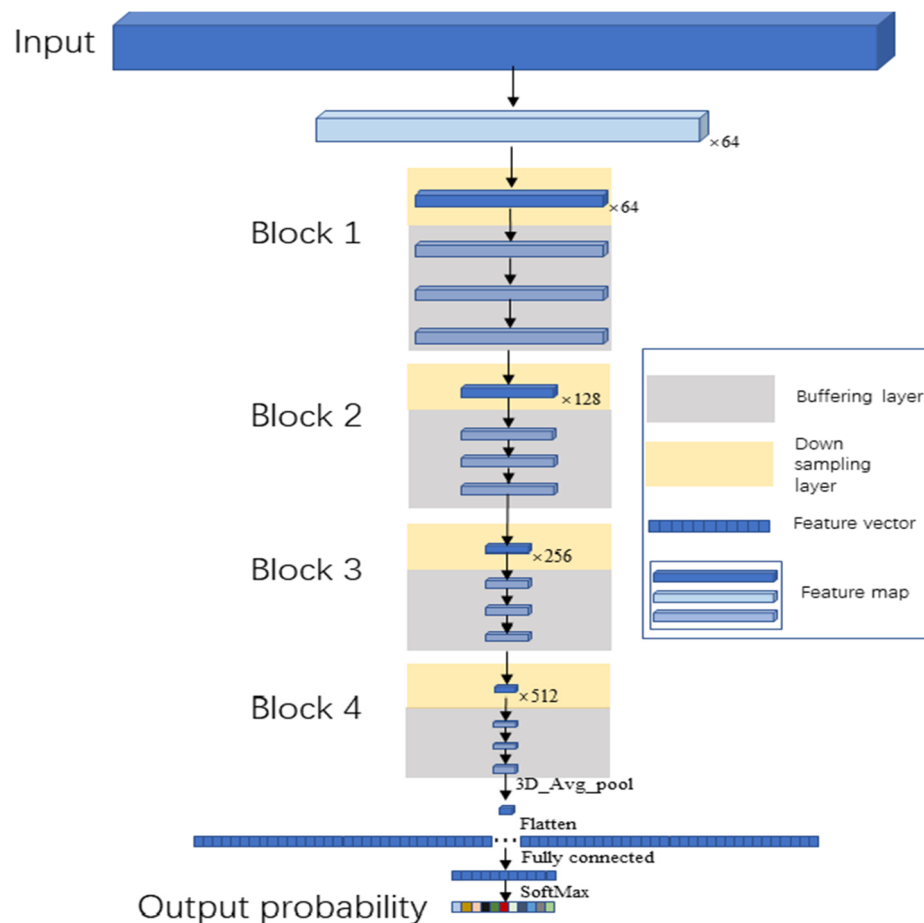


Figure 3. The architecture of the Buffer Net and the data flow. The input was HMI images at a size of $200 \times 25 \times 25$. Firstly, it was down sampled with the output as 64 feature maps, whose sizes were $98 \times 25 \times 25$. Then, they were input to the four blocks to down sample and buffer. The size of the output feature maps for the four blocks were $49 \times 13 \times 13$, $25 \times 7 \times 7$, $13 \times 4 \times 4$ and $7 \times 2 \times 2$, respectively, and the number of output feature maps for the four blocks were 64, 128, 256 and 512, respectively. The feature map output by the fourth block was flattened into a vector after a 3D average pooling, which was then input to the fully connected layer. The final output was the probability for each bacterial genus by SoftMax. For accelerating the training process, every 3D convolutional layer was followed by a 3D Batch Normalization layer.

2.5. Development Language and Training Details

In this study, Python was adopted to develop the Buffer Net for identifying bacterial genera. In addition, C# was adopted to develop the interactive software for the AI-assisted system. The Buffer Net was integrated into this software to accomplish the specimen digitization, image analysis and bacteria identification.

The batch size was set to 128. All images were resized to 25×25 for accelerating the algorithm training before inputting the model. For training, we set the learning rate to 0.0001, the momentum to 0.9 and the epoch to 30. The optimizer used in the paper is the stochastic gradient descent (SGD). At each epoch, the training samples were reshuffled.

2.6. System Integration

The whole bacterial identified system was established by integrating the HMI system, the interactive software with a 3D-CNN-based algorithm (Buffer Net) component, data preprocessing and genus identification. The advantage of the whole system is that it could perform all of these steps in a minute, with 30 s for imaging and another 30 s for data preprocessing and genus identification.

To achieve rapid data storage, a USB 3.0 interface was adopted to connect with the computer, which coincided with the data digitization. Once it was completed, the pseudo color image was automatically produced and shown on the screen. The identification was achieved by clicking on the identification buffer of the user. All information about the cases could be exported by the system in the form of a PDF (Figure 2c). The appearance and components of this system are shown in Figure 2.

Furthermore, to facilitate the subsequent model upgrades, an interface in the software was set up to quickly invoke the model. The relevant model developed by Python, the weight file obtained from training and packaging as a file in the form of .exe were all packed for easy use. The subsequent upgrades would be achieved by replacing the .exe file without much modification to the code.

2.7. Evaluation Metrics

In this study, the confusion matrix and overall accuracy were applied to evaluate the performance of the system for differentiating these 11 genera. In the confusion matrix, the ordinate represents the true genus of each category, and the abscissa represents the genus predicted by algorithms. Diagonal elements represent the prediction accuracy of each genus, and nondiagonal elements represent the percentage that the genus on the ordinate was misjudged as the category on the abscissa. The overall accuracy represents the total precision of the system for differentiating all these 11 genera.

The calculation method for the elements in the (I, J) coordinates in the confusion matrix is as follows:

$$Precision_{ij} = \frac{P_{ij}}{\sum_{k=1}^{11} P_{ik}} \times 100\%, \quad (1 \leq i \leq 11, 1 \leq j \leq 11).$$

$Precision_{ij}$ represents the percentage where i genera were misidentified as j genera. P_{ij} represents the number of genera i identified as j genera, and P_{ik} represents the number of genera i identified as k genera.

In order to evaluate the overall accuracy of the system, we also introduced the overall accuracy, whose calculation formula is as follows:

$$Accuracy = \frac{\sum_{i=1}^{11} P_{ii}}{\sum_{j=1}^{11} \sum_{k=1}^{11} P_{jk}} \times 100\% = \frac{\sum_{i=1}^{11} P_{ii}}{\sum_{j=1}^{11} n_j} \times 100\% = \frac{\sum_{i=1}^{11} P_{ii}}{N} \times 100\%$$

P_{ii} represents the number of genera i bacteria correctly classified, and P_{ij} represents the number of j genera predicted as genera k . Notably, j could be equal to k . n_j represents the total number of bacteria diagnosed as j genera, and N represents the total number of bacteria of all the genera.

3. Results

3.1. Hyperspectral Microscopic Images

Over 130,000 hyperspectral images of 11 common infectious pathogens were collected with cautiousness (Stenotrophomonas, Escherichia, Morganella, Burkholderia, Serratia, Pseudomonas, Acinetobacter, Klebsiella, Proteus, Staphylococcus and Enterococcus) for individual bacteria through hyperspectral microscopic imaging technology.

The datacube of a bacterium (Pseudomonas) is presented in Figure 4. For ease of observation, several wavelength images of the datacube are presented separately. Moreover, the mean spectrum of a single bacterium is also visualized.

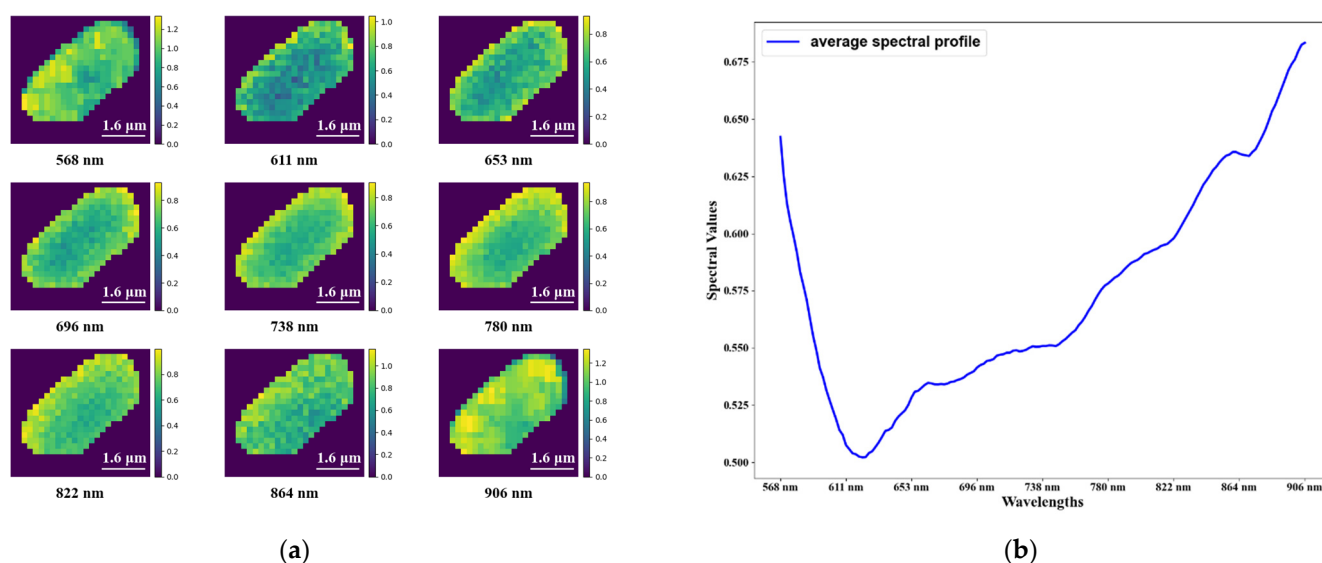


Figure 4. (a) The heatmap of a bacterium (*Pseudomonas*) on several wavelengths. The brightness of the color in the images indicates the intensity of the energy. The background in the heatmap is filled with zero. (b) The average spectral profiles for a single bacterium.

3.2. Classification Performance of the AI-Assisted System

As shown in Table 2, the overall accuracy of the Buffer Net for identifying a bacterial genus was 94.9%, which was the best and higher than 3D-ResNet (92.3%) [28]. The 1D-CNN was the worst, with an accuracy of only 82.6%. Meanwhile, the accuracy for 2D-CNN was moderate, whose accuracy was 91.3%. Without buffering layers, Buffer Net got a low accuracy, 88.4%. All of these results were acquired on the test dataset with over 40,000 hyperspectral images.

Table 2. The comparison of the differentiation accuracy for all samples.

Algorithm	1D-CNN	2D-CNN	3D-ResNet	Buffer Net (Without)	Buffer Net (With)
Accuracy	82.6	91.3	92.3	88.4	94.9

Red represents the highest performance. **With/without** means Buffer Net with/without buffering layers.

3.3. The Differentiation Speed of Our AI-Assisted HMI System

As for the speed, it took only 30 s for our HMI system to differentiate genera on microscopic hyperspectral images at the cellular level.

4. Discussion

An AI-assisted pathogen detection system, with designed HMI and Buffer Net, was developed for rapid, accurate and low-price automatic identification in this paper. Before us, Seo et al. and Rui et al. used HMI to classify food-borne bacteria [17–20]. Nevertheless, our method was better than the previous studies in terms of data or algorithm.

In terms of data, the wavelength range of our hyperspectral data is 440–1023 nm, and the spectral resolution is 2.1 nm, with a total of 277 bands. In contrast, the wavelength range in these researches is 450–800 nm, with a spectral resolution of 4 nm and a total of 89 bands. In addition, the maximum number of images in those studies was no more than 5000, far less than the number (>130,000) we collected. Concerning algorithms, machine learning, 1D-CNN, 2D-CNN and LSTM were utilized to extract features that would disconnect the strong spatial–spectral correlation in datacubes. Here, we proposed a new method based on 3D-CNN, called Buffer Net, to train the model directly from datacubes.

As for the accuracies of differentiating bacterial genera for 1D-CNN, 2D-CNN, 3D-ResNet and Buffer Net, it was no surprise that 1D-CNN has the lowest accuracy (detailed

in Figure 5). In contrast, Buffer Net had the highest, because only spectral information was adopted for the 1D-CNN with spatial features discarded. Likewise, it was explainable that the accuracy of 2D-CNN was higher than that of 1D-CNN and lower than that of 3D-ResNet, because the images were considered conventional images with numerous channels, which enabled the simple usage of the spectral profiles with spatial features adopted, while 3D-ResNet could extract the features from the spatial and spectral dimensions precisely.

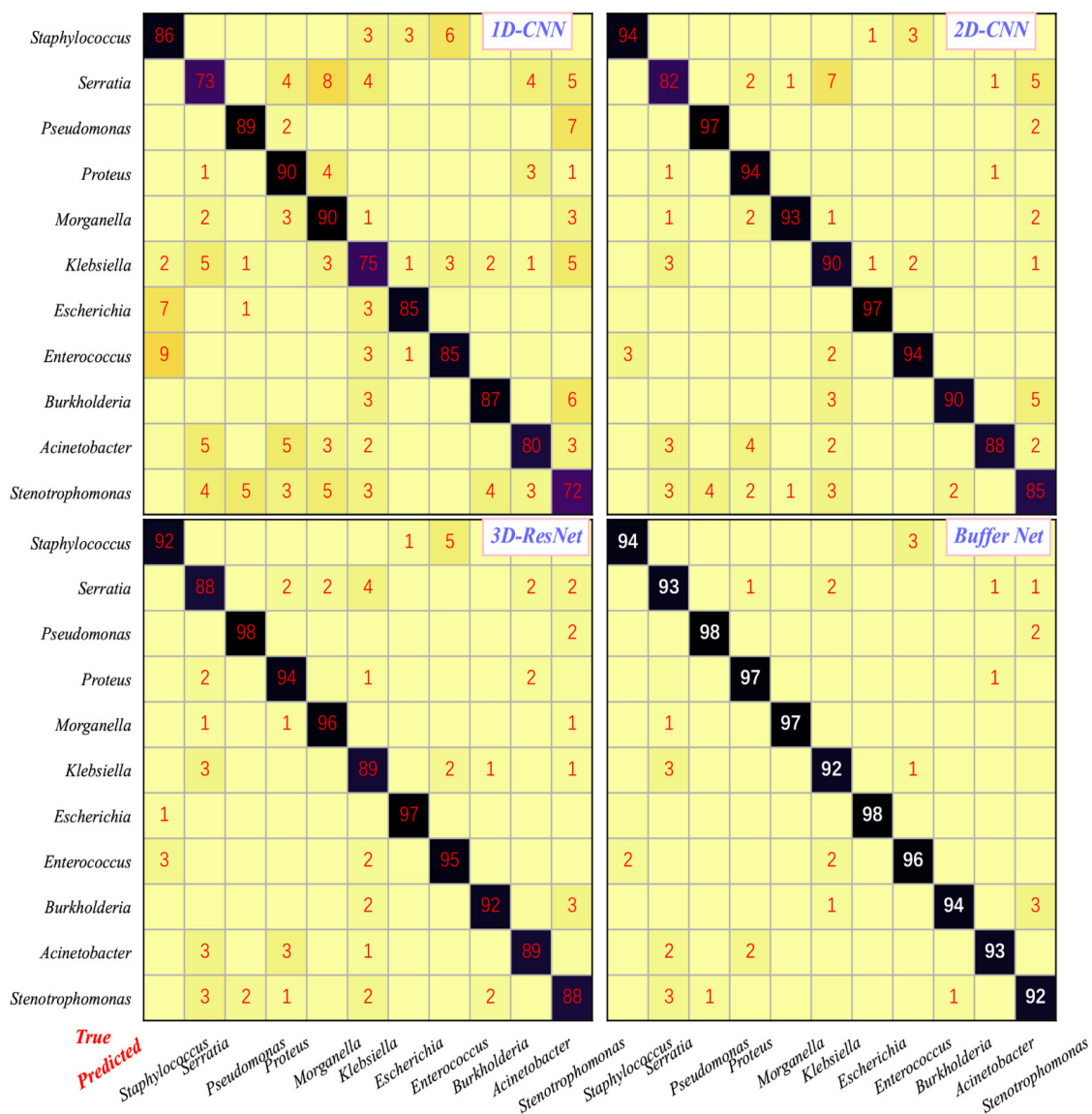


Figure 5. The confusion matrix of differentiating genera of Buffer Net, 1D-CNN, 2D-CNN and 3D-ResNet. The x-axis represents the genera predicted by our system, while the y-axis represents the underlying facts. Entries (i, j) indicate the percentage of samples that actually belong to class i that are judged to be of class j ; entries along the diagonal indicate the discrimination accuracy of the model we developed for each genus, and off-diagonal entries indicate the misclassification accuracy. The algorithm name was presented in the top right corner of each subplot. The highest accuracy for identifying each bacterial genus was marked by bold white among these four algorithms. Accuracies with less than 1% were not presented in the confusion matrix. All classification probabilities shown in red. The probability on the diagonal represents the correct rate. In particular, white represents the highest accuracy of the four algorithms for each genus. The remaining probability that is not diagonally represents the error rate.

Noticeably, our proposed algorithm, Buffer Net, achieved the highest accuracy of 94.9%, because it was designed to learn higher-level features. It is noticeable that the differentiation precisions of Buffer Net for all genera were over 92%, with the highest precision of 98% for *Pseudomonas* and *Escherichia*. Therefore, the requirement for clinical practice was met. Buffer Net held the highest precision for identifying 11 genera compared with the others.

ResNet is a classical deep architecture that achieves a good performance in many tasks [28]. However, in our experiments, the overall accuracy of 3D-ResNet (92.3%) is 2.6% lower than Buffer Net (94.9%). Since both ResNet and Buffer Net in this paper use 3D-CNN as filters, the difference in performance can only be caused by the difference of architecture. To further demonstrate the effectiveness of the Buffer Net architecture, we practiced an ablation experiment by removing all buffering layers from Buffer Net and obtained an accuracy of 88.4% (Table 2). The accuracy of Buffer Net without the buffering layers is even less accurate than 2D-CNN (91.3%), indicating that the frequent down sampling limits the feature representation capability, which is enhanced by the buffering layers.

Compared with conventional methods in routine practice, our AI-assisted system allows for automatic and rapid bacterial identification. Since it identifies the genus of individual bacteria, it does not require any culture of the pathogen, as long as the individual bacteria can be isolated. In scenarios where bacterial isolation techniques are unavailable, the system can significantly reduce the cultivating time of the bacteria. Compared to expensive PCR and MALDI TOF, our method can obtain high precision while keeping the costs under control, even lower than the morphological examination. In addition, since the operation is simple and the system automates the identification, there is no expertise required for the operators. Therefore, this can reduce the cost of the clinical diagnosis for training physicians and their redundant workloads.

Recently, time-lapse coherent imaging [29], three-dimensional quantitative phase imaging [30] and holographic microscopy [31] have been used for bacterium identification. Nevertheless, time-lapse coherent imaging based on colonies is time-consuming. Three-dimensional quantitative phase imaging or the holographic microscopy identifies pathogens with a refractive index (RI). HMI uses spectral and morphological information to classify pathogen genera.

This study also had some limitations. Firstly, the precision of those algorithms mentioned above for identifying these four genera (*Serratia*, *Klebsiella*, *Acinetobacter* and *Stenotrophomonas*) was relatively low, and this may be due to the similarities between the four genera. Further improvements will be implemented to optimize the precisions. In addition, we focused on bacterial genera identification at the cellular level in this study, which could be further subdivided into species and subspecies. Species or subspecies identification based on datacubes of pathogens will be explored in subsequent studies, and these are what we are doing. Lastly, although over 130,000 images with 11 genera were adopted to develop the algorithm Buffer Net, far more than 11 genera need to be identified in routine clinical practice. Therefore, the number and diversity of the training and test data need to be augmented further.

5. Conclusions

In this study, we developed an AI-assisted system to automatically identify bacterial genera by combining the designed HMI system and Buffer Net. Utilizing the HMI system, a large-scale dataset where bacteria datacubes had a wider wavelength range and a finer spectral resolution was built to train Buffer Net for identifying 11 genera. The 3D-CNN and the architecture with buffering layers and down sampling layers contributed to the highest accuracy (94.9%) of Buffer Net compared with the other algorithms. The AI-assisted system has great potential in rapidly identifying bacteria, because it can shorten or even eliminate the cultivation time. After obtaining the microscopic datacube of an individual bacterium, it only takes 30 s to classify a pathogen.

Author Contributions: Conceptualization, C.T., Y.T., B.H. and Z.Z.; Data curation, B.H. and Z.Z.; Formal analysis, C.T. and J.D.; Funding acquisition, B.H., Z.Z. and K.D.; Investigation, C.T., Y.T., J.W., Z.Z. and M.Y.; Methodology, C.T., J.D. and Z.Z.; Project administration, B.H., Z.Z. and K.D.; Resources, B.H., K.D. and M.Y.; Software, C.T., J.D. and J.W.; Supervision, J.D., B.H., Z.Z. and K.D.; Validation, C.T.; Visualization, C.T. and J.W.; Writing—original draft, C.T., J.D. and Y.T. and Writing—review and editing, C.T., J.D. and Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This study was granted by the project of “Research on automatic hyperspectral pathology diagnosis technology” (Project No. Y855W11213) supported by the Key Laboratory Foundation of Chinese Academy of Sciences and the project of “Research on microscopic hyperspectral imaging technology” (Project No. Y839S11) supported by the Key Laboratory of Biomedical Spectroscopy of Xi’an.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data will be made available by the corresponding author upon reasonable request. The code is at <https://github.com/chengtao/Buffer-Net>; The trained weights is at <https://drive.google.com/file/d/123AYGNOAIKJTWLv6ugSbEZ4RwSK8nN3f/view?usp=sharing>.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Gan, Y.; Li, C.; Peng, X.; Wu, S.; Li, Y.; Tan, J.P.; Yang, Y.Y.; Yuan, P.; Ding, X. Fight bacteria with bacteria: Bacterial membrane vesicles as vaccines and delivery nanocarriers against bacterial infections. *Nanomed. Nanotechnol. Biol. Med.* **2021**, *35*, 102398. [[CrossRef](#)] [[PubMed](#)]
- Rangel-Vega, A.; Bernstein, L.R.; Mandujano Tinoco, E.-A.; García-Contreras, S.-J.; García-Contreras, R. Drug repurposing as an alternative for the treatment of recalcitrant bacterial infections. *Front. Microbiol.* **2015**, *6*, 282. [[CrossRef](#)] [[PubMed](#)]
- Sadarangani, M. Protection against invasive infections in children caused by encapsulated bacteria. *Front. Immunol.* **2018**, *9*, 2674. [[CrossRef](#)] [[PubMed](#)]
- van Elsland, D.; Neeffjes, J. Bacterial infections and cancer. *EMBO Rep.* **2018**, *19*, e46632. [[CrossRef](#)]
- Versalovic, J. *Manual of Clinical Microbiology*; American Society for Microbiology Press: Washington, DC, USA, 2011; Volume 1.
- Engelmann, I.; Alidjinou, E.K.; Ogiez, J.; Pagneux, Q.; Miloudi, S.; Benhalima, I.; Ouafi, M.; Sane, F.; Hober, D.; Roussel, A. Preanalytical issues and cycle threshold values in SARS-CoV-2 real-time RT-PCR testing: Should test results include these? *ACS Omega* **2021**, *6*, 6528–6536. [[CrossRef](#)]
- Alseekh, S.; Aharoni, A.; Brotman, Y.; Contrepolis, K.; D’Auria, J.; Ewald, J.; Ewald, J.C.; Fraser, P.D.; Giavalisco, P.; Hall, R.D. Mass spectrometry-based metabolomics: A guide for annotation, quantification and best reporting practices. *Nat. Methods* **2021**, *18*, 747–756. [[CrossRef](#)]
- Rave, A.; Kuss, A.; Peil, G.; Ladeira, S.; Villarreal, J.; Nascente, P. Biochemical identification techniques and antibiotic susceptibility profile of lipolytic ambient bacteria from effluents. *Braz. J. Biol.* **2018**, *79*, 555–565. [[CrossRef](#)]
- Park, H.-T.; Ha, S.; Park, H.-E.; Shim, S.; Hur, T.Y.; Yoo, H.S. Comparative analysis of serological tests and fecal detection in the diagnosis of *Mycobacterium avium* subspecies *paratuberculosis* infection. *Korean J. Vet. Res.* **2020**, *60*, 117–122. [[CrossRef](#)]
- Roux-Dalvai, F.; Gotti, C.; Leclercq, M.; Hélie, M.-C.; Boissinot, M.; Arrey, T.N.; Daully, C.; Fournier, F.; Kelly, I.; Marcoux, J. Fast and Accurate Bacterial Species Identification in Urine Specimens Using LC-MS/MS Mass Spectrometry and Machine Learning. *Mol. Cell. Proteom.* **2019**, *18*, 2492–2505. [[CrossRef](#)]
- Leekha, S.; Terrell, C.L.; Edson, R.S. General principles of antimicrobial therapy. In *Mayo Clinic Proceedings*; Elsevier: Amsterdam, The Netherlands, 2011; pp. 156–167.
- Alexandrakis, D.; Downey, G.; Scannell, A.G. Detection and identification of bacteria in an isolated system with near-infrared spectroscopy and multivariate analysis. *J. Agric. Food Chem.* **2008**, *56*, 3431–3437. [[CrossRef](#)]
- Yoon, S.-C.; Lawrence, K.; Siragusa, G.; Line, J.; Park, B.; Feldner, P. Hyperspectral reflectance imaging for detecting a foodborne pathogen: *Campylobacter*. *Trans. ASABE* **2009**, *52*, 651–662. [[CrossRef](#)]
- Windham, W.R.; Yoon, S.-C.; Ladely, S.R.; Heitschmidt, J.W.; Lawrence, K.C.; Park, B.; Narrang, N.; Cray, W.C. The effect of regions of interest and spectral pre-processing on the detection of non-0157 Shiga-toxin producing *Escherichia coli* serogroups on agar media by hyperspectral imaging. *J. Near Infrared Spectrosc.* **2012**, *20*, 547–558. [[CrossRef](#)]
- Yoon, S.-C.; Windham, W.R.; Ladely, S.R.; Heitschmidt, J.W.; Lawrence, K.C.; Park, B.; Narang, N.; Cray, W.C. Hyperspectral imaging for differentiating colonies of non-0157 Shiga-toxin producing *Escherichia coli* (STEC) serogroups on spread plates of pure cultures. *J. Near Infrared Spectrosc.* **2013**, *21*, 81–95. [[CrossRef](#)]

16. Kammies, T.-L.; Manley, M.; Gouws, P.A.; Williams, P.J. Differentiation of foodborne bacteria using NIR hyperspectral imaging and multivariate data analysis. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 9305–9320. [[CrossRef](#)] [[PubMed](#)]
17. Seo, Y.; Park, B.; Hinton, A.; Yoon, S.-C.; Lawrence, K.C. Identification of Staphylococcus species with hyperspectral microscope imaging and classification algorithms. *J. Food Meas. Charact.* **2016**, *10*, 253–263. [[CrossRef](#)]
18. Kang, R.; Park, B.; Eady, M.; Ouyang, Q.; Chen, K. Single-cell classification of foodborne pathogens using hyperspectral microscope imaging coupled with deep learning frameworks. *Sens. Actuators B Chem.* **2020**, *309*, 127789. [[CrossRef](#)]
19. Kang, R.; Park, B.; Eady, M.; Ouyang, Q.; Chen, K. Classification of foodborne bacteria using hyperspectral microscope imaging technology coupled with convolutional neural networks. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 3157–3166. [[CrossRef](#)]
20. Kang, R.; Park, B.; Ouyang, Q.; Ren, N. Rapid identification of foodborne bacteria with hyperspectral microscopic imaging and artificial intelligence classification algorithms. *Food Control* **2021**, *130*, 108379. [[CrossRef](#)]
21. Seibert, J.A.; Boone, J.M.; Lindfors, K.K. Flat-field correction technique for digital detectors. In *Medical Imaging 1998: Physics of Medical Imaging*; SPIE: Bellingham, WA, USA, 1998; pp. 348–354.
22. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [[CrossRef](#)]
23. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Processing Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
24. Tian, S.; Wang, S.; Xu, H. Early detection of freezing damage in oranges by online Vis/NIR transmission coupled with diameter correction method and deep 1D-CNN. *Comput. Electron. Agric.* **2022**, *193*, 106638. [[CrossRef](#)]
25. Hsieh, T.-H.; Kiang, J.-F. Comparison of CNN algorithms on hyperspectral image classification in agricultural lands. *Sensors* **2020**, *20*, 1734. [[CrossRef](#)] [[PubMed](#)]
26. Ji, S.; Xu, W.; Yang, M.; Yu, K. 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 221–231. [[CrossRef](#)] [[PubMed](#)]
27. Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 740–755.
28. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
29. Wang, H.; Ceylan Koydemir, H.; Qiu, Y.; Bai, B.; Zhang, Y.; Jin, Y.; Tok, S.; Yilmaz, E.C.; Gumustekin, E.; Rivenson, Y. Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning. *Light Sci. Appl.* **2020**, *9*, 118. [[CrossRef](#)]
30. Kim, G.; Ahn, D.; Kang, M.; Jo, Y.; Ryu, D.; Kim, H.; Song, J.; Ryu, J.S.; Choi, G.; Chung, H.J. Rapid and label-free identification of individual bacterial pathogens exploiting three-dimensional quantitative phase imaging and deep learning. *BioRxiv* **2019**. [[CrossRef](#)]
31. Jo, Y.; Park, S.; Jung, J.; Yoon, J.; Joo, H.; Kim, M.-H.; Kang, S.-J.; Choi, M.C.; Lee, S.Y.; Park, Y. Holographic deep learning for rapid optical screening of anthrax spores. *Sci. Adv.* **2017**, *3*, e1700606. [[CrossRef](#)] [[PubMed](#)]