

Rapid Detection of Bacteria Using Raman Spectroscopy and Deep Learning

Kaitlyn Kukula, Denzel Farmer, Jesse Duran, Nishatul Majid, Christie Chatterley, Jeff Jessing, Member, IEEE, and Yiyan Li, Member, IEEE

Department of Physics and Engineering, Fort Lewis College, Durango, CO, USA

Abstract— Bacteria identification can be a time-consuming process. Machine learning algorithms that use deep convolutional neural networks (CNNs) provide a promising alternative. Here, we present a deep learning based approach paired with Raman spectroscopy to rapidly and accurately detect the identity of a bacteria class. We propose a simple 4-layer CNN architecture and use a 30-class bacteria isolate dataset for training and testing. We achieve an identification accuracy of around 86% with identification speeds close to real-time. This optical/biological detection method is promising for applications in the detection of microbes in liquid biopsies and concentrated environmental liquid samples, where fast and accurate detection is crucial. This study uses a recently published dataset of Raman spectra from bacteria samples and an improved CNN model built with TensorFlow. Results show improved identification accuracy and reduced network complexity.

Keywords—bacteria, Raman spectroscopy, deep learning, rapid detection.

I. INTRODUCTION

As bacteria identification plays a role in a variety of fields (including medical, forensic, and environmental), there is a great need for faster bacteria identification methods [1, 2]. Traditional methods of identifying bacterial strains are time-consuming processes that often require an incubation period of 24-48 hours [3]. Raman spectroscopy is a valuable and versatile tool that can identify bacteria genotypes [4-6]. The Raman spectra act as nondestructive and unique fingerprints for their samples [6, 7]. The principle behind the technology is the Raman effect, where a sample exposed to a laser will scatter light at a frequency different from that of the incident light (these are known as vibrational signatures) [6]. The process of generating spectra is considerably simpler and faster than standard culturing and allows for higher throughput, fewer preparation materials (which risk error), and non-specialists to gather high-quality data [6, 8].

Chi-sing Ho et al. published a dataset with 60,000 Raman spectra from 30 bacterial isolates, representative of most infections seen in intensive care units worldwide [4]. This reduces the barrier of entry for building faster and more accurate neural networks that can outperform traditional methods. The Raman spectroscopy dataset stores each spectrum as a one-dimensional array of 1000 points representing its vibrational signature.

Several neural network techniques have been applied to bacteria identification, the most successful being convolutional and residual networks [4]. Complex architectures promise to improve on traditional spectral data analysis. These networks can be extremely deep, some with upwards of 150 layers. While this depth can produce a highly-fit network, those networks run the risk of overfitting data which adds unnecessary complexity and slows processing [9, 10]. Developing a simplified network to achieve accurate and generalizable results, with minimal layers, is one of the ambitions of our approach.

In our experiment, we used Ho's dataset to train a CNN that is fast and accurate with the objective of being able to learn to identify new bacterial strains. Our CNN could cut down the number of generations, allowing for an overall faster network while maintaining an improved accuracy. Many industries, including forensic analysis, medical, and environmental can benefit from this technique.

II. METHODS

For this experiment, we prepared two neural networks. The first one was a recreation of the neural architecture published previously [4]. This network was composed of 26 layers and included an initial convolution layer and four residual blocks, each with six convolutional layers per residual block and a final dense classification layer. Residual skip connections allow the model to skip residual blocks that proved too deep for the network. This network was built for modularity and allowed us to adjust a wide variety of hyperparameters, including the structure of the network.

Through experimentation with the modular network, we developed a second network which examined the role of network depth on accuracy. This simpler convolutional neural network model was constructed of three convolutional layers, a pooling layer, and a final dense classifier layer for a total network depth of four layers. This model structure is visualized in Figure 1. An Adam optimizer, learning rate of 0.001, batch size of 100, and a dropout of 0.65 is used. Convolutional filters for each convolutional layer are respectively 10, 25, and 25.

Both networks used pretraining and finetuning stages to optimize the performance of the model. The networks were pretrained on the large 60,000 sample database, and finetuned on a much smaller 3,000 sample dataset, which included samples more similar to the testing dataset. They each ran over

10 epochs for pretraining and 10 epochs for finetuning. Average accuracies for both models were calculated over 10 trials. The datasets were preprocessed using a rolling mean.

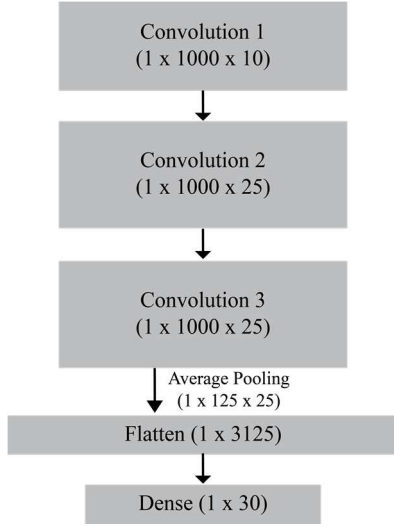


Figure 1. Model diagram of a 4-layer network, featuring three convolution layers and a final dense layer.

III. RESULTS

Sample Raman signals obtained from the bacteria dataset are plotted in Figure 2. These signals are obtained from the Chi Sing Ho et al. dataset [4]. Signals are processed using a rolling mean, which is plotted over unprocessed signals to show how a rolling mean accentuates identifying characteristics of spectra. Plotting of signals shows that different strains have similar characteristics in their Raman spectra.

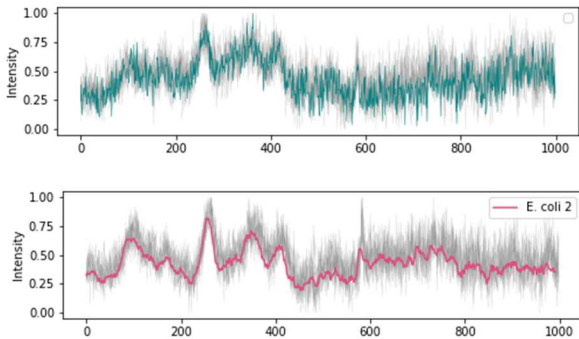


Figure 2. Sample data signals plotted. Signals represent one commonly misidentified spectra, *E. coli 2* (bottom) and what it is commonly misidentified as *K. pneumoniae 1* (top).

The average accuracy achieved by the 26-layer residual network model with pretraining was $85.29 \pm 0.47\%$ with an average runtime of 264.47 seconds per trial (on a benchmark Google Colaboratory environment). The average accuracy achieved by the 4-layer convolutional network with pretraining was $85.82 \pm 0.78\%$ with an average runtime of 59.22 seconds per trial in the same environment. Without pretraining, the average accuracy was $81.06 \pm 0.83\%$ for the 26-layer network

and $80.23 \pm 1.22\%$ for the 4-layer network. Accuracies of the models with and without pretraining are visualized in Figure 3, along with the average runtimes. With pretraining, the 4-layer network is about 22.4 times faster in detection than the 26-layer network (Table 1).

The accuracy and loss of the 4-layer convolution network's pretraining and finetuning models can be seen in Figure 3. The accuracy of the model converges after 10 epochs of pretraining and 5 epochs of finetuning.

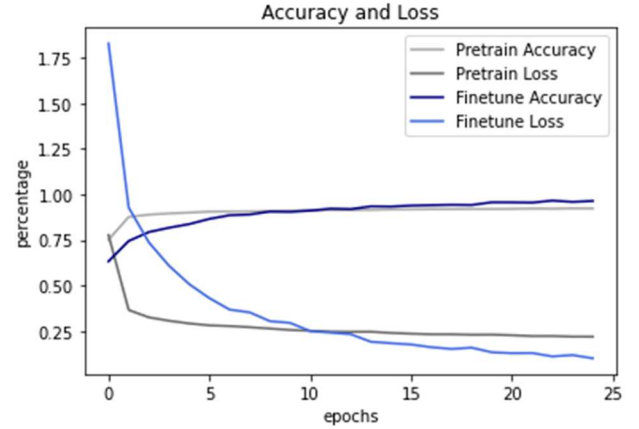


Figure 3. Plot representing the accuracy and loss function of the pretraining and finetuning models of the 4-layer convolutional network. The models are run over 20 epochs for pretraining and finetuning for one trial.

The critical learning rate hyperparameter was evaluated and a parametric study of the hyperparameter value is displayed in Figure 4. The learning rate was optimized to be 0.0001 for pretraining and finetuning on the 26-layer model and 0.001 for the 4-layer model.

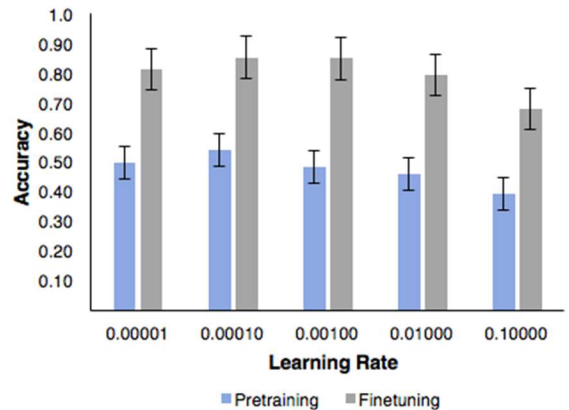


Figure 4. Analysis of the effect of learning rate on the accuracy of the 26-layer model after finetuning, assessed over 10 trials with standard deviation represented by error bars.

Major elements of the model were considered and tuned, including the number of training epochs and how many pretraining samples the network required for maximum accuracy. Hyperparameters, runtime, and accuracy achieved

were analyzed to optimize the model. The number of epochs begins to converge after 10 epochs in pretraining and 5 epochs in finetuning for the 4-layer model; thus, only 10 epochs are needed to optimize the identification accuracy and limit the total runtime. As expected, the total runtime for the 4-layer network is significantly lower than that of the 26-layer network and achieves similar identification accuracies (Table 1). The number of epochs is consistent between the two models, so network depth plays a key role in runtime and should be minimized to gain efficiency.

Table 1. Comparison of model runtime to complete 10 trials and accuracy achieved by 26-layer network that replicates the residual network developed by Chi-Sing Ho et al. and the 4-layer convolutional network, both with and without pretraining. Accuracy was tested over 10 trials with the standard deviation of accuracy provided for each model.

	26 Layer Network		4 Layer Network	
	With Pretraining	Without Pretraining	With Pretraining	Without Pretraining
Runtime [sec]	2644.71	150.04	592.26	32.71
Average Accuracy	85.29%	81.06%	85.82%	80.23%
Standard Deviation	0.0047	0.0083	0.0078	0.012

The normalized confusion matrix, seen in Figure 5, displays the identification accuracy of the 30 bacterial classes across the wall. Identifications of less than 0.5% are not displayed. The true bacteria class is on the y-axis while the class predicted by the model is on the x-axis. This matrix was obtained from one trial of the 4-layer convolutional network.

IV. DISCUSSION

The 4-layer network provides accurate identifications of bacterial isolates with significantly reduced runtime. Most

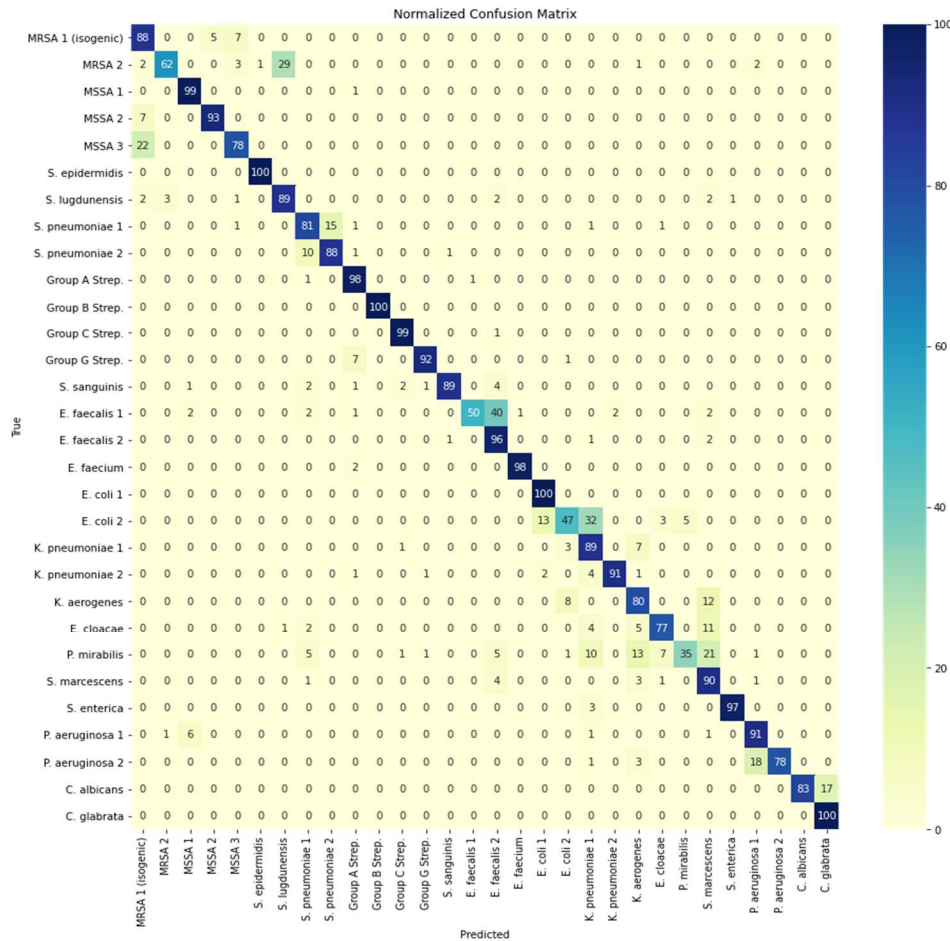


Figure 5. Normalized confusion matrix with entries along the diagonal representing maximum identification accuracies of the 30 bacterial isolates used in this study. Obtained from one trial on the 4 layer network, which achieved an overall accuracy of 85.27% for the trial.

notably is the *E. coli* 1 and *E. coli* 2 (Figure 5) identifications, which will be the focus of future study due to the common use of *E. coli* as an indicator of water quality, specifically fecal contamination. Although the 26-layer model has comparable accuracy and standard deviation for identification of *E. coli* 1 and *E. coli* 2, training the network is time-consuming and the model contains more layers than are required for the simple task of identifying basic Raman signals. Both models have a difficult time differentiating between similar bacteria isolates, such as *P. mirabilis* and *E. faecalis* 1. This is likely due to their Raman signals having similar basic components, such as a peak or trough in a similar position, as seen in Figure 5. Figure 2 shows *E. coli* 2 and *K. pneumoniae* 1. *E. coli* 2 is commonly misidentified as *K. pneumoniae* 1, and more processing of the Raman signals or further hyperparameter tuning may be needed in order to consistently identify *E. coli* strains accurately for use in industries such as environmental monitoring, where *E. coli* is a common indicator.

Further data preprocessing may aid in reducing noise and extracting unique signal features, which will aid the model in differentiating between similar bacteria isolates in the future. The loss function of both pretraining and finetuning on the 4-layer network is promising as it quickly drops to around 0.10 and converges. Our model can quickly identify key components of the bacteria isolate signals and continue to maintain high rates of identification after each round of optimization.

Pretraining the neural network is helpful as it extracts basic features of the signal and saves data that will be accessed in finetuning, leading to overall higher identification accuracy. Pretraining the 26- and 4-layer networks lead to about a 4% increase in overall identification accuracy, as seen in Table 1. Further, pretraining the network decreases the standard deviation of the identifications and will thus lead to a more precise model. Pretraining does account for much of the runtime of both models, but a single pretrained network could be finetuned for multiple datasets. The convolutional layers extract general features during pretraining, and only specific details are updated during finetuning.

The learning rate was also evaluated, as outlined in Figure 4. This hyperparameter is vital to the success of the neural network as it controls the convergence rate of the network. For values greater than the optimal value, the network likely was subject to an unstable or divergent learning process; for values less than the optimal value, the network was unable to process the data quickly, resulting in a slightly longer runtime. In the future, CNN's will be applied to a variety of other Raman signals sourced from aquatic-based bacteria. Further preprocessing and normalizing of the dataset will likely reduce misidentification errors.

V. ACKNOWLEDGMENTS

This study was supported by undergraduate research internships from NSF PREM (#1827847), NSF STROBE (#1548924), NSF REU (#1757953), EPA P3 (#SU83988001, this publication was developed under Assistance Agreement No. SU83988001 awarded by the U.S. Environmental

Protection Agency to Yiyan Li. It has not been formally reviewed by EPA. The views expressed in this document are solely those of Yiyan Li and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.), and NIH-Maximizing Access to Research Careers Undergraduate Student Training in Academic Research (MARC U*STAR) award number T34GM092711-FLC. We also thank Dr. Chi-sing Ho's insightful discussion.

VI. REFERENCES

- [1] M. G. Madden and A. G. Ryder, "Machine learning methods for quantitative analysis of Raman spectroscopy data," in *Opto-Ireland 2002: Optics and Photonics Technologies and Applications*, 2003, pp. 1130-1139.
- [2] X. Lu, H. M. Al-Qadiri, M. Lin, and B. A. Rasco, "Application of mid-infrared and Raman spectroscopy to the study of bacteria," *Food and Bioprocess Technology*, vol. 4, pp. 919-935, 2011.
- [3] K. Qu, F. Guo, X. Liu, Y. Lin, and Q. Zou, "Application of machine learning in microbiology," *Frontiers in Microbiology*, vol. 10, p. 827, 2019.
- [4] C.-S. Ho, N. Jean, C. A. Hogan, L. Blackmon, S. S. Jeffrey, M. Holodniy, et al., "Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning," *Nature communications*, vol. 10, pp. 1-8, 2019.
- [5] S. Pahlow, S. Meisel, D. Cialla-May, K. Weber, P. Rösch, and J. Popp, "Isolation and identification of bacteria by means of Raman spectroscopy," *Advanced drug delivery reviews*, vol. 89, pp. 105-120, 2015.
- [6] T. Huser and J. Chan, "Raman spectroscopy for physiological investigations of tissues and cells," *Advanced drug delivery reviews*, vol. 89, pp. 57-70, 2015.
- [7] L. Gómez-Mascaraque, K. Kilcawley, D. Hennessy, J. Tobin, and T. O'Callaghan, "Raman spectroscopy: A rapid method to assess the effects of pasture feeding on the nutritional quality of butter," *Journal of Dairy Science*, vol. 103, pp. 8721-8731, 2020.
- [8] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, et al., "Using Raman spectroscopy to characterize biological materials," *Nature protocols*, vol. 11, pp. 664-687, 2016.
- [9] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6848-6856.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.