Supporting Information

Rapid SERS identification of methicillin-susceptible and methicillin-resistant *Staphylococcus aureus* via aptamer recognition and deep learning

Shu Wang[a,b], Hao Dong[a,b], Wanzhu Shen[d], Yong Yang[a,b], Zhigang Li[a], Yong Liu[a,b], Chongwen Wang[d*], Bing Gu[c*] , and Long Zhang[a,b*]

[a] Hefei Institute of Physical Science, Chinese Academy of Sciences, Hefei 230036, PR China.

[b.] University of Science and Technology of China, Hefei 230036, PR China.

[c] Laboratory Medicine, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510000, PR China

[d.] Anhui Agricultural University, Hefei 230036, PR China.

[*]Corresponding author. *Email address:*

zhanglong@aiofm.ac.cn (Z. Long),

gb20031129@163.com (B. Gu),

wangchongwen1987@126.com (C. Wang).

S1. Experimental section

S1.1 Materials and chemicals

Sodium borohydride ($NaBH_4$), $AgNO_3$, $HAuCl_4 \cdot 4H_2O$, trisodium citrate, and sodium chloride (NaCl) were purchased from Shanghai Chemical Reagent Co., Ltd. (China). The clinical isolates of *S. aureus*, *E. coli*, *Acinetobacter baumannii* (*A. baumannii*), *P. aeruginosa*, *Streptococcus pneumoniae* (*S. pneumoniae*), *Staphylococcus epidermidis* (*S. epidermidis*), *L. monocytogenes* and *Salmonella typhimurium* (*S. typhimurium*) were obtained from the Affiliated Hospital of Xuzhou Medical University. The experiment procedure was approved by the Ethics Committee of the Institute of the Affiliated Hospital of Xuzhou Medical University. The aptamers for *S. aureus* (aptamer$_{Sa}$) and *E. coli* (aptamer$_E$) were synthesized by Sangon Biotech (Shanghai).

Aptamer$_{Sa}$:

5'-
GCAATGGTACTTCCACTTAGGTCGAGGTTAGTTTGTCTTGCTGGCGCATCC
ACTGAGCGCAAAAGTGCACGCTACTTTGCTAA-3';

Aptamer$_E$: 5'-
GCAATGGTACGGTACTTCCTCGGCACGTTCTCAGTAGCGCTCGCTGGTCAT
CCCACAGCTACGTCAAAAGTGCACGCTACTTTGCTA-3'

The *S. aureus* and *E. coli* aptamers used in this study was isolated by Shao et al. and Kim et al. respectively, by using bacterial cell-SELEX.

S.1.2 Instruments

Transmission electron microscope (TEM) images were obtained using a Hitachi H-7650 TEM at an accelerating voltage of 80 kV. Dynamic light scattering with a Brookhaven Zeta PALS instrument was used to measure the zeta potential of the prepared nanoparticles and bacteria samples The UV-Vis spectra were measured using a Shimadzu 2600 spectrometer. The Raman spectrum was recorded with a portable Raman system (B&W Tek, i-Raman Plus BWS465–785H spectrometer). All bacteria samples were excited by a 785-nm laser with a power of 25 mW and a total acquisition time of 20 s for each SERS spectrum. Five spectra from each sample were collected and averaged to ensure signal reproducibility.

S1.3 Preparation of bacterial sample

The ordinary plate counting method was used to determine bacterial concentrations. The experiment procedure was performed according to the previous publication. In brief, *S. aureus*, *and E. coli* were cultured at 37 °C for 5 h in Luria–Bertani (LB) medium. Then, 0.1 mL of the bacterial culture was diluted with LB medium for 1 × 10^5 times, coated onto the agar plates, and cultured at 37 °C overnight. Finally, the number of colony-forming units (CFUs) was counted and thus the orignal bacteria concentration can be calculated.

S1.4 SERS detection protocol

A total of 0.1 mL of MSSA and MRSA were incubated with 200 nM aptamer for 20 min. Then, the *S. aureus*-aptamer complexes were centrifuged at 4000 rpm for 4 min

and the supernatant was discarded. The precipitate was redispersed in 100 μL of $AgNO_3$ solution (10 mM), and the mixture was intensely vortexed for 1 min. Subsequently, 100 μL of $NaBH_4$ solution was added into the tube, and mixture was vortexed intensely for another 1 min. The formed bacteria-aptamer@Ag complexes were collected by centrifugation, and then dropped on a Si substrate for the measurement of their SERS spectra.

S1.5 Spectrum measurements

For each species, bacteria samples were prepared across 25 strains chosen at random in 30 with the concentration of 10^7 cells/mL, and we measured Raman spectra at 40 sites randomly of a solution from each strain. In general, 1000 MRSA and MSSA SERS spectra were measured for model training, respectively.

By the same approach, the independent test dataset was measured at bacteria solution from the 10 strains bacteria (5 strains of MRSA and 5 strains of MSSA) with the concentration of 10^5 cells/mL. 40 SERS spectra were measured from one bacterial strain, thus total 400 spectra (200 of MRSA and 200 of MSSA) were used for prediction at one cell density.

To assess the repeatability of the proposed method, 10 different batches of MSSA and MRSA samples (~10^7 cells/mL) were tested. The relative standard deviation (RSD) values of the SERS intensities at 730 cm^{-1} were calculated by the following equation:

$$\mathrm{RSD} = \frac{Standard\ deviation}{Mean\ SERS\ intensities}$$

S1.6 CNN architecture and training

Fig. 4 displayed the shallow CNN model used in this work, it is mainly composed by two convolutional layers and one fully connected layer (Dense in Fig.4). At the first beginning, the input data shape of spectra is 664 × 1 (550 cm^{-1}-1800 cm^{-1}). In the first and second convolutional layers, 32 and 128 kernels(filters) with the mask size of 9 × 1 were used to do convolution operation with the input data, respectively. Followed every convolutional layer, the outputs of the convolutional layer were transmitted into

an activation function of “Relu”, for the introduce of nonlinear factor. The input spectra with data shape of 664 × 1 was reshaped as 648 × 128 after two convolutional operation and two activation operation. In flatten layer, the input data with shape of 648 × 128 was merged and reorganized as a 1D vector with shape of 82944 × 1, and was output into fully-connected layer (dense). The fully-connected layer is the last layer and it generally connect the predict tags with each unit of the past 1D vector using weighted matrices and a “sigmoid” activation function for binary classification task. In model compiling, we utilized a 10-fold cross-validation in model training, 90% data were used to generate parameters of 1D filters and the weighted matrices, then the last 10% data were verified by CNN with current parameters. When model iterates, the parameters were auto adjusted to purchase the minimum predict error until average accuracy is stable. The “binary_crossentropy” loss function is used for the binary task herein, and we use “sigmoid” optimizer with learning rate 0.001 and batch size 32, the number iterates at the reference data for 200 epochs. We also experimented with a 30 layers depth network based on residual architectures, but found that the shallow architecture performed best in binary task.

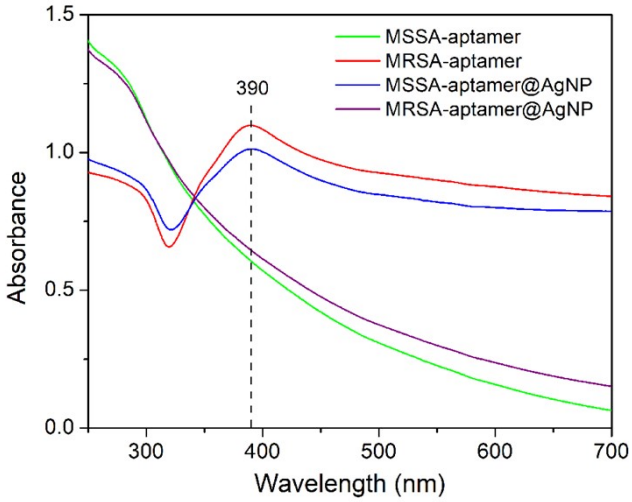

Fig. S1 UV-vis spectra of the formed MSSA–aptamer, MRSA–aptamer, MSSA–aptamer@AgNP, and MRSA–aptamer@AgNP.

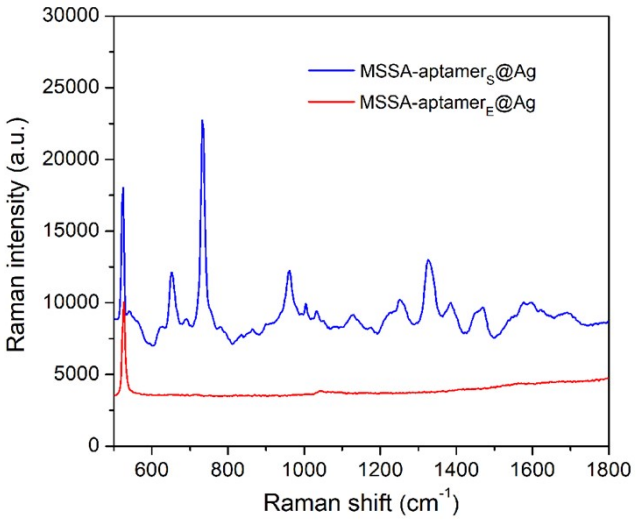

Fig. S2 SERS spectra of MSSA–aptamer$_S$@AgNP (blue line) and MSSA–aptamer$_E$@AgNP (red line).

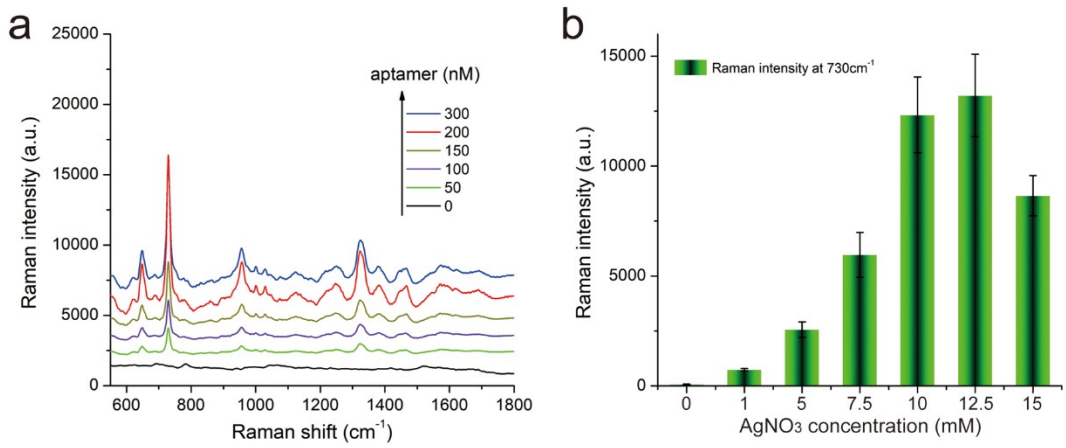

Fig. S3 Optimization of (a) aptamer concentration and (b) $AgNO_3$ concentration for *S. aureus* detection.

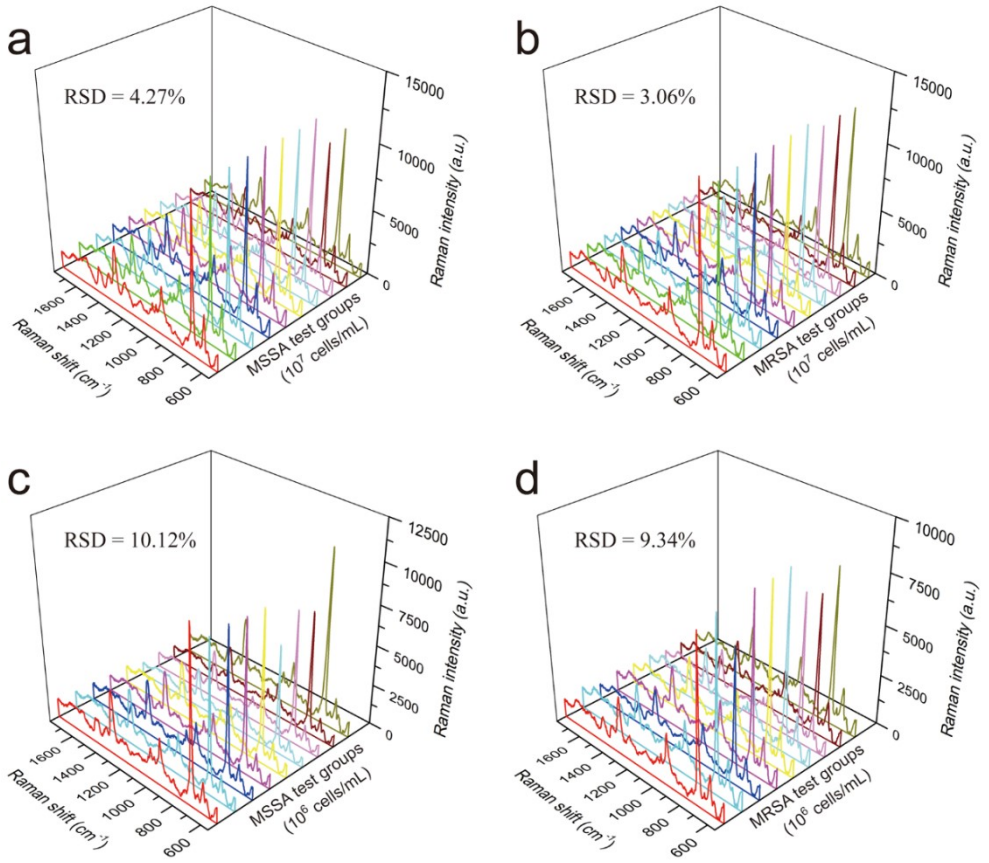

Fig. S4 SERS spectra of ten different batches bacteria: (a) MSSA-aptamer@AgNP (10^7 cells/mL); (b) MRSA-aptamer@AgNP (10^7 cells/mL), (c) MSSA-aptamer@AgNP (10^6 cells/mL) and (d) MRSA-aptamer@AgNP (10^6 cells/mL).

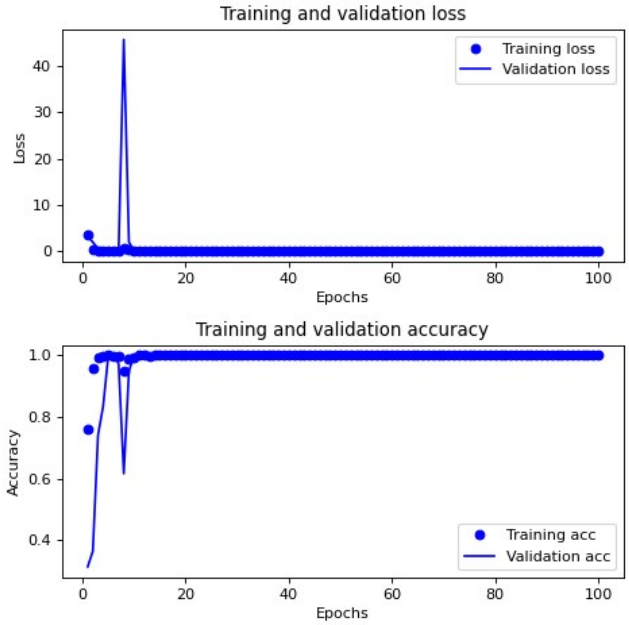

Fig. S5 The loss and accuracy of training data and validation data in the process of model iteration, the point represents training data, the line represents validation data.

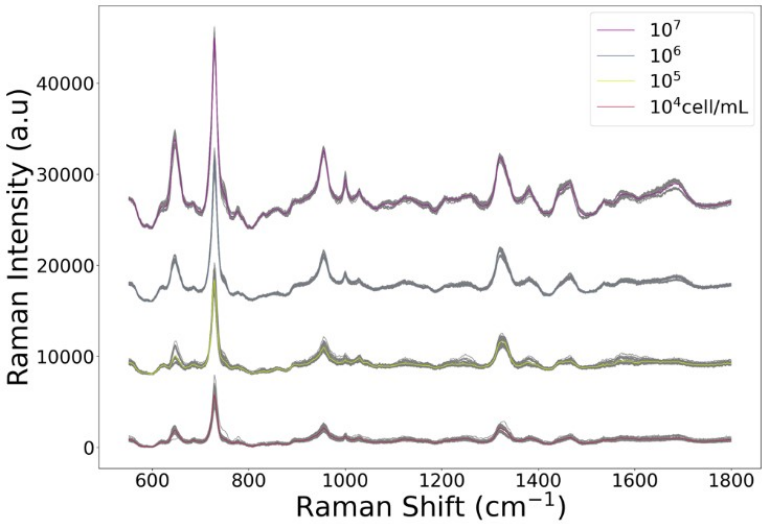

Fig. S6 SERS spectra take from the *S. aureus*–aptamer@Ag complexes with different concentrations of *S. aureus* (10^7-10^4 cells/mL).

Table S1 Identification performance of the CNN classifier on Raman spectra of different bacterial concentrations (10^7-10^4 cells/mL).

	Concentration			
	10^4 cells/mL	10^5 cells/mL	10^6 cells/mL	10^7 cells/mL
Identyfication accuracy for MRSA	98	100	100	100
Identyfication accuracy for MSSA	95	100	100	100
Average accuracies (%)	96.5	100	100	100

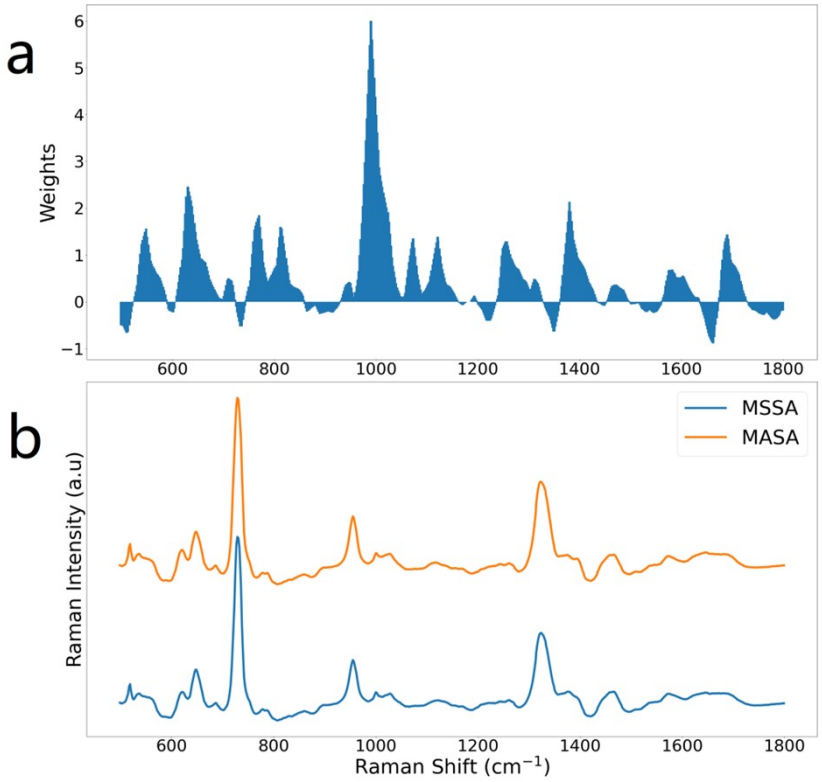

Fig. S7 Contribution of each Raman bands to the correct classification.

In the classification processing of CNN based classifier, each element in the original data is associated with the final decision through a weight. Gradient-weighted class activation mapping (Grad-CAM) is an algorithm that can be used to output all weights as a contribution matrix (***Int J Comput Vision***, 2020, 128, 336-359). In the flatten layer of the constructed CNN model, the initial input spectra of 664 × 1 was transformed into a 648 × 128 matrix through convolution operation, and each element in this matrix has matching a weight from the contribution matrix for the final decision of classification. In our study, we have accumulated the contribution matrix row by row to get a 648 × 1 vector, then reshaped the new vector as the same size of

original spectra (664 × 1) to characterization the contribution of each spectral bands to the correct classification. The spectral weight vector was shown in Fig. S7. The value corresponding to the abscissa reflect the importance of this band to distinguish MRSA from MSSA. It can be observed that the bands of 955-1040 cm^{-1} make the largest contribution. According to Table 1, these signals are from the molecular vibration of proteins and cyanides.