(Received 4 August 2025, Accepted 13 September 2025, Scheduled 14 September 2025)

# Raman and FTIR Fingerprint Spectra of Blood and Bronchoalveolar Lavage Fluid for AI-Based Classification of Severe Pneumonia

Sailing He[1, 2, 3, *], Jialun Li[1], Anqi Yang[1], Chenghui Wang[4, 2], Chuan Zhang[1, 2], Xinyue Li[1, 3], Ke Cui[2, 5], Youzu Xu[6], Julian Evans[1], and Yinghe Xu[2, 5]

[1] *Centre for Optical and Electromagnetic Research, National Engineering Research Center for Optical Instruments College of Optical Science and Engineering, Zhejiang University, Hangzhou 310058, China*
[2] *Zhejiang Engineering Research Center for Intelligent Medical Imaging, Sensing and Non-invasive Rapid Testing Taizhou Hospital of Zhejiang Province, Wenzhou Medical University, Linhai 317000, China*
[3] *MedEngInfo Collaborative Research Center Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou 318000, China*
[4] *Department of Laboratory Medicine Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai 317000, China*
[5] *Department of Critical Care Medicine Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou 318000, China*
[6] *Department of Respiratory and Critical Care Medicine Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou 318000, China*

**ABSTRACT:** Severe pneumonia poses a significant threat to public health. Delayed diagnosis is a core challenge in treatment. This study uses two rapid, low-cost spectroscopic fingerprinting techniques — Raman spectroscopy and attenuated total reflectance Fourier transform infrared (ATR-FTIR) absorption spectroscopy — to analyze biofluids such as blood and bronchoalveolar lavage fluid (BALF). In contrast to our earlier work which combined infrared spectra with clinical biochemical test results, this paper focuses solely on the spectral data to validate a fast and label-free diagnostic method. We used a spectral transformer network (STNetwork) to perform AI-based classification of severe pneumonia from the spectral fingerprints of blood and BALF. While both modalities are effective, FTIR spectroscopy exhibits superior diagnostic precision (97.78% test accuracy) and stability (SD $<$ 0.0139) for blood samples. BALF offers a unique window into the local lung microenvironment, and both metabolomic analysis and spectral fingerprint classification were performed. The classification results for BALF Raman spectra (enhanced with surface-enhanced Raman spectroscopy) gave a training accuracy of $96.71\% \pm 1.86\%$ and a testing accuracy of $90.62\% \pm 3.95\%$, better than the classification results for BALF FTIR spectra. The present study provides a reliable technical foundation for developing rapid and high-accuracy screening solutions for severe pneumonia.

## 1. INTRODUCTION

Severe pneumonia [1], a critical respiratory disease with high global morbidity and mortality rates, poses a serious threat to public health. According to World Health Organization (WHO) data, over 3 million people worldwide die from severe pneumonia annually, with mortality rates for severe community-acquired pneumonia reaching 30%–50% [2]. Approximately 60% of patients are diagnosed only after progressing to moderate or advanced stages, and delayed diagnosis is a key factor affecting the prognosis [3]. The key challenge in the precise diagnosis and treatment of severe pneumonia lies in the difficulty of early warning, as current assessment systems are not timely enough.

Conventional methods fall short in capturing the dynamic progression of the disease. Early clinical symptoms (such as body temperature) and conventional infection markers (such as white blood cell count) lack specificity. Conventional pulmonary imaging of severe pneumonia lacks specificity in early diagnosis, though lung CT images are useful in identifying severe bacterial infection (or even mild bacterial infection) [4, 5]. While deep-seated pathophysiological changes (such as microcirculatory perfusion disorders and mitochondrial energy metabolism dysfunction) offer valuable diagnostic information, they are difficult to monitor in real time.

Recent advancements in genomics, transcriptomics, proteomics, and metabolomics ('omics' research) have provided insights into the underlying mechanisms of severe pneumonia. These techniques have identified pathogen spectra, host immune responses, functional protein changes, and metabolic disturbance [6]. However, the practical application of 'omics' research in a clinical setting is limited by several factors: the need for expensive equipment, high costs, and long turnaround times — often taking more than a week to yield results. Thus, they are unsuitable for continuous monitoring or rapid diagnosis, especially for critically ill patients who require timely intervention.

* Corresponding author: Sailing He (sailing@zju.edu.cn).

Currently, diagnostic and classification methods based on clinical manifestations, traditional microbiological testing, and single biomarkers (such as PCT and CRP) do not accurately capture the dynamic progression of diseases or provide early warnings of deterioration risks [7]. They also fall short in guiding personalized treatment decisions, leading some patients to miss the optimal window for intervention.

To address these clinical challenges, we turn to spectroscopic fingerprinting technology for rapid diagnosis of severe pneumonia. Spectroscopic techniques, including Raman spectroscopy and FTIR spectroscopy, can provide a rapid, comprehensive biochemical snapshot of biological samples by analyzing their unique vibrational spectra. These "fingerprints" contain rich information about the molecular composition of the sample, reflecting changes in proteins, lipids, nucleic acids, and carbohydrates associated with disease states. These methods are label-free, low-cost, and have the potential to provide results in minutes rather than days.

**Raman Spectroscopy** is a vibrational spectroscopy technique that provides information about molecular vibrations, rotations, and other low-frequency modes. It measures the inelastic scattering of monochromatic light, typically from a laser. When light interacts with a molecule, most photons are elastically scattered (Rayleigh scattering), but a small fraction undergoes inelastic scattering (Raman scattering) where the photons gain or lose energy. This energy shift corresponds to the vibrational modes of the molecule, creating a unique spectral fingerprint. Raman spectroscopy has advantages such as high spatial resolution, minimal water interference, and the ability to analyze samples in their native state. However, Raman signals are inherently weak and can be difficult to detect, often requiring signal enhancement techniques like Surface-enhanced Raman spectroscopy (SERS) for certain applications.

**FTIR Spectroscopy** measures the absorption of infrared radiation by a sample. Specific chemical bonds in the sample absorb infrared radiation at characteristic frequencies, causing them to vibrate. The resulting absorption spectrum provides a fingerprint of the molecular composition of the sample. ATR-FTIR is an excellent technique for analyzing liquid samples such as blood and BALF, since it requires only a small volume of sample and minimal preparation. ATR (Attenuated Total Reflection) simplifies FTIR analysis by enabling the direct measurement of samples through the interaction of an evanescent wave with the sample's surface, requiring little to no preparation. The primary advantage of FTIR spectroscopy is the high signal intensity and excellent signal-to-noise ratio, so it is a robust and reliable technique. It can be developed into a micro-FTIR imaging (see, e.g., [8]).

In a prior study [9], we demonstrated the benefits of a multi-modal approach for the early screening of acute myeloid leukemia (AML) by combining serum FTIR spectra with standard clinical blood biochemical test results. We showed that this multi-modal classification system, which used a multi-modality spectral transformer network (MSTNetwork), achieved higher accuracy and sensitivity than methods relying on a single modality, and improved the sensitivity by 12% (compared with using only biochemical indicators). The present work aims to validate the diagnostic potential of spectral fingerprinting alone, without any external clinical data, for the classification of severe pneumonia. This is a crucial step towards developing a standalone and rapid diagnostic tool that can be used for continuous monitoring in critical care settings, a capability that traditional clinical tests cannot provide. The ability to measure biological samples in short time is essential for the precise management of severe pneumonia. Unlike biochemical tests, which provide a static snapshot at the time of the blood draw and cannot be performed continuously, spectral measurements can be acquired in real-time, offering a dynamic view of a patient's condition.

## 2. MEASUREMENT AND METHODOLOGY

### 2.1. Sample Cohorts and Clinical Data

The study used two separate cohorts for spectroscopic analysis:

- **Blood Cohort:** This cohort consisted of **194 subjects**, including **40 patients with severe pneumonia** and a control group of healthy individuals. All samples were collected and processed in accordance with ethical guidelines and institutional review board approval. One pneumonia sample from this cohort could only be measured by infrared absorption spectroscopy due to insufficient residual volume for Raman spectroscopy.

- **Bronchoalveolar Lavage Fluid (BALF) Cohort:** This cohort was composed of 105 patients who had undergone a BALF procedure. The group included **69 patients with severe pneumonia** and a control group of **36 patients with non-severe pneumonia** The purpose of this cohort was to compare the spectral fingerprints between the two groups.

All samples were processed and stored according to standard protocols to preserve their integrity.

### 2.2. Spectroscopic Measurements

- **Raman Spectroscopy:** We used a self-developed Raman imaging system for data acquisition (see [10] and [11]). For BALF samples, which are known to have a weak intrinsic signal, we utilized our home-made SERS (Surface-enhanced Raman spectroscopy) chips (see, e.g., [11]) to enhance the Raman signals. For blood samples, which have a stronger intrinsic signal, we did not use SERS enhancement

- **ATR-FTIR Spectroscopy:** The infrared absorption spectra were measured using an ATR-FTIR spectrometer. This method is particularly well-suited for biological fluids like blood and BALF due to its minimal sample volume requirements and ease of use. All the liquid samples (blood or BALF) were directly measured without the time-consuming freeze-drying process. Before each measurement, the detection area was cleaned with anhydrous ethanol to prevent contamination. Subsequently, ultrapure

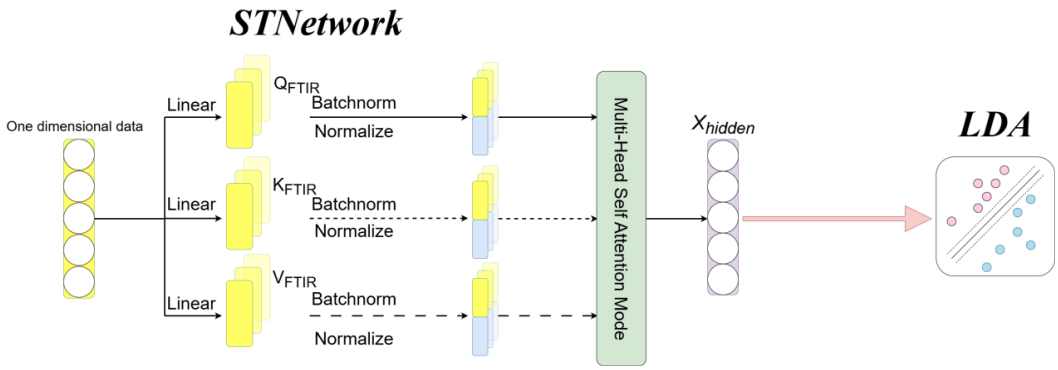


**FIGURE 1**. The structure of Spectral Transformer Network (STNetwork).

water (as a reference) was drawn into an appropriate volume using a pipette and placed on the diamond crystal attachment. The pressure head evenly adhered the liquid to the crystal as a background for scanning. After the scanning was completed, the detection area was cleaned with anhydrous ethanol, and the liquid sample was tested using the same procedure (like that for the ultrapure water). At least five samples were taken and tested for each liquid sample, and the average value was used.

### 2.3. AI-Based Classification

The Raman/FTIR spectra were treated as long one-dimensional vectors with high-dimensional spectral data (each wavenumber can be regarded as a feature variable). To effectively analyze the high-dimensional spectral data, we employ a deep learning architecture adapted from our previous work [9], which we call the Spectral Transformer Network (STNetwork). This network is specifically designed to first extract hierarchical features and then model the global dependencies within those features. As shown in Fig. 1, the architecture consists of two main stages: (1) a feature extraction backbone using a one-dimensional Residual U-Net (ResUNet), and (2) a feature refinement stage using a Multi-Head Self-Attention mechanism.

The feature vector produced by the ResUNet backbone is then passed to a Multi-Head Self-Attention module for further refinement. In this single-modality context, the attention mechanism functions as a self-attention layer, enabling the model to understand the global dependencies and intricate relationships within the extracted feature set. The process begins by linearly projecting the input feature vector into three distinct representations: the Query (Q), Key (K), and Value (V). Since all three are derived from the same source, the model is essentially learning to attend to itself. To improve training stability, BatchNorm1d and L2 normalization are applied to these projections.

The core of the module computes scalar dot-product attention scores, which quantify the compatibility between different parts of the feature vector. The model can dynamically weight the importance of each feature in the context of all other features. The "multi-head" design allows this process to be performed in parallel across several representation subspaces, enabling the network to jointly capture various types of feature interactions. The outputs from all heads are concatenated and passed through a final linear layer to produce a single, contextually-aware, and highly discriminative feature vector. This refined vector serves as the final output of the network, ready for downstream tasks like classification.

## 3. RESULTS

### 3.1. Classification of Severe Pneumonia Using Fingerprint Spectra of Blood Samples

The results of the AI-based classification of blood samples from the first cohort (severe pneumonia patients versus healthy controls) show clear differences in both Raman and FTIR.

#### *3.1.1. Patients versus Healthy Controls*

The average Raman spectra of blood samples from severe pneumonia patients and healthy controls are shown in Figure 2. The plot shows clear differences in the average Raman signals between the two groups. The shaded area represents the standard deviation ($\pm 1\sigma$).

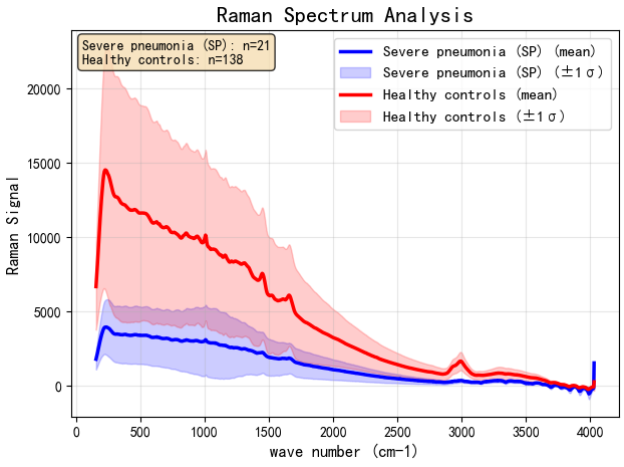


**FIGURE 2**. Raman spectra of blood samples from severe pneumonia (SP).

The confusion matrices of the AI-based classification results for blood Raman spectra are shown in Figure 3. After five independent validation trials, the model demonstrated excellent classification capabilities with a training set accuracy of

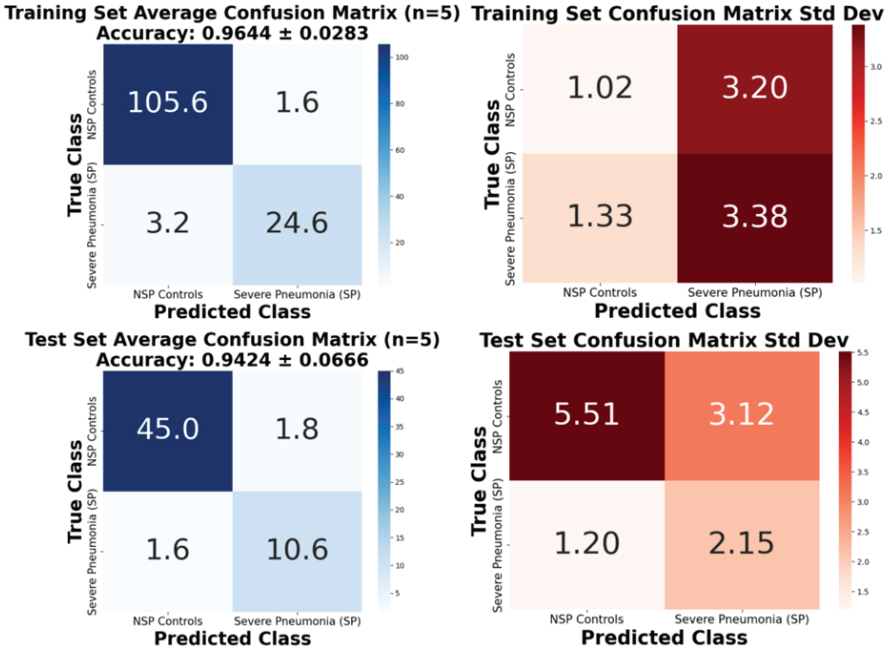


**FIGURE 3**. AI-based classification results using blood Raman spectra of severe pneumonia (SP) patients and healthy controls.

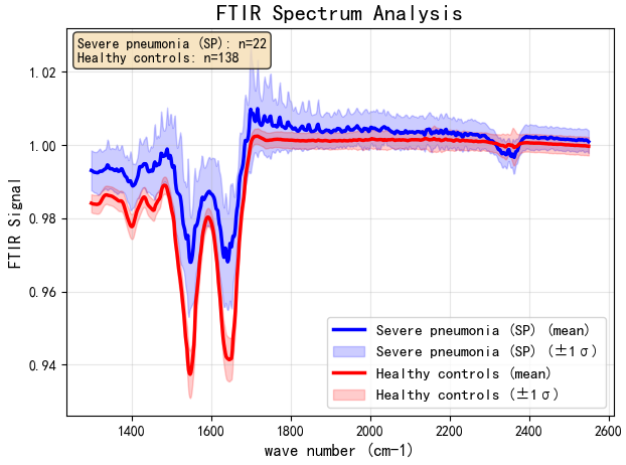


**FIGURE 4**. Infrared absorption spectral fingerprints of blood samples from severe pneumonia (SP) patients versus healthy controls.

96.44%±2.83% and a testing set accuracy of 94.24%±6.66%. All accuracy differentials between training and testing sets were within 2%, confirming model generalizability without overfitting.

Figure 4 shows the average infrared absorption spectral fingerprints of blood samples from severe pneumonia patients and healthy controls. The spectra show differences in peak shape and intensity. The shaded areas represent the standard deviation ($\pm 1\sigma$).

The AI-based classification results for blood infrared absorption spectra are shown in Figure 5. This modality demonstrated superior performance: training set accuracy was 99.20% ± 1.24% and testing set accuracy was 97.78%±1.39%. The accuracy differential was within 2%, and the data showed superior experimental stability (SD $< 0.02$ vs Raman's $< 0.1$).

## 3.2. Bronchoalveolar Lavage Fluid (BALF) Analysis

### *3.2.1. Metabolomic Analysis of BALF*

While clinical scores and inflammatory markers in blood are widely used, their ability to capture the full biological complexity of the pneumonia is limited. BALF offers a unique window into the local lung microenvironment, as its components directly reflect the biological state of infection, inflammation, and tissue damage. In-depth molecular analysis of BALF has the potential to reveal specific pathophysiological mechanisms of severe pneumonia and inform the development of more precise diagnostic tools and treatments.

Here we identify and characterize the unique metabolite profiles in the BALF of patients with severe pneumonia (SP) and non-severe pneumonia (NSP) using non-targeted metabolomics. Due to the high cost of multi-omics analysis, a large sample size is generally not used for such studies. We selected BALF samples from eight patients with severe pneumonia and eight with non-severe pneumonia at Taizhou Hospital for non-targeted metabolomics analysis. Although the sample size is small, it is sufficient for an exploratory study designed to generate new hypotheses and identify potential biomarkers, as it can capture significant differential signals between disease states. As shown in Figure 6, all detected metabolites in positive ion (POS; panel A) and negative ion (NEG; panel B) modes were primarily classified as lipids and lipid-like molecules, organic acids and derivatives, benzenoids, organic oxygen compounds, and others. Principal component analysis (PCA) revealed minimal variability in metabolomic profiles within each group (severe pneumonia, SP; non-severe pneumonia, NSP), while significant differences were observed between the two groups.

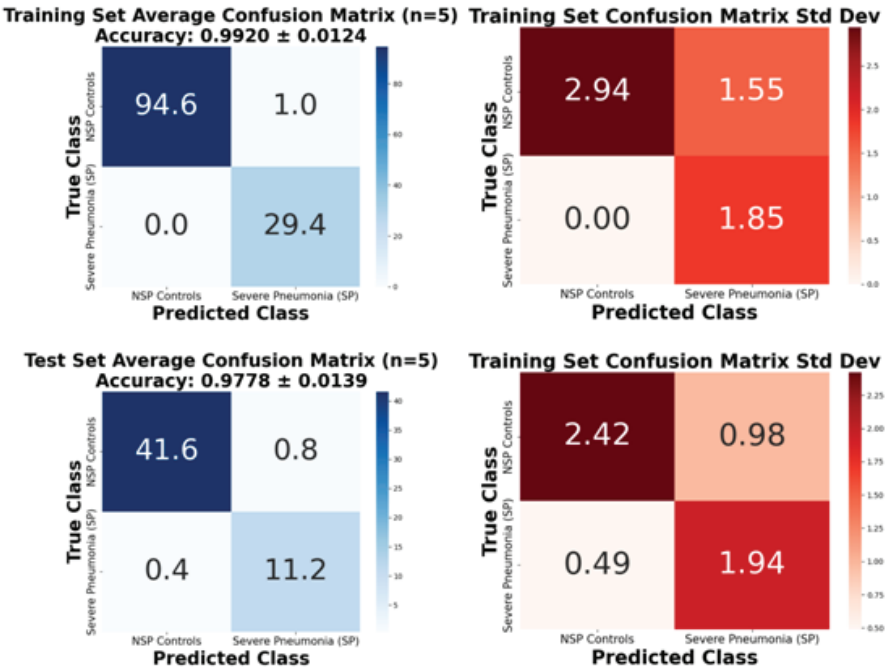


**FIGURE 5**. Detection and classification results using blood infrared absorption spectra from pneumonia patients (SP) versus healthy controls.

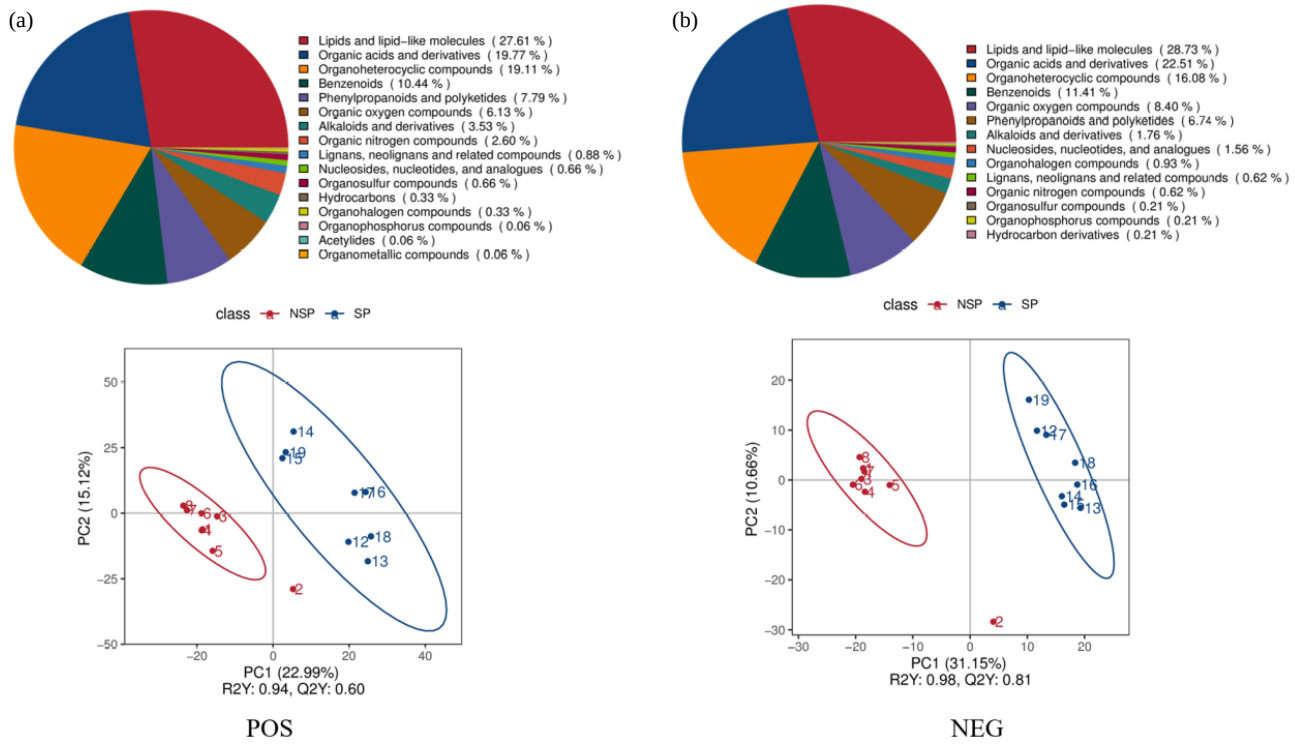


**FIGURE 6**. Metabolite classification overview and principal component analysis (PCA) of BALF from patients with severe pneumonia versus non-severe pneumonia controls. The axes represent the PC1 and PC2 scores of the samples. Samples with positive scores exhibit above-average levels of the metabolite combinations constituting PC1 or PC2, whereas samples with negative scores display below-average levels of those metabolite combinations. A larger absolute value of the PC1 or PC2 score indicates a greater deviation of the sample's metabolic profile from the overall average metabolic profile. The percentage associated with PC1 (PC2) represents the proportion of the total variance in the original data that can be explained by PC1 (PC2) R2Y represents the interpretability of the model, indicating the percentage of information in the model that can explain the predefined categorical variable Y. The closer R2Y is to 1, the more information the model can explain regarding the classification groups, and the greater the difference between the two groups. Q2Y indicates the model's explanatory power when predicting new (unseen) data. A higher Q2Y value suggests better predictive performance and greater reliability of the model.

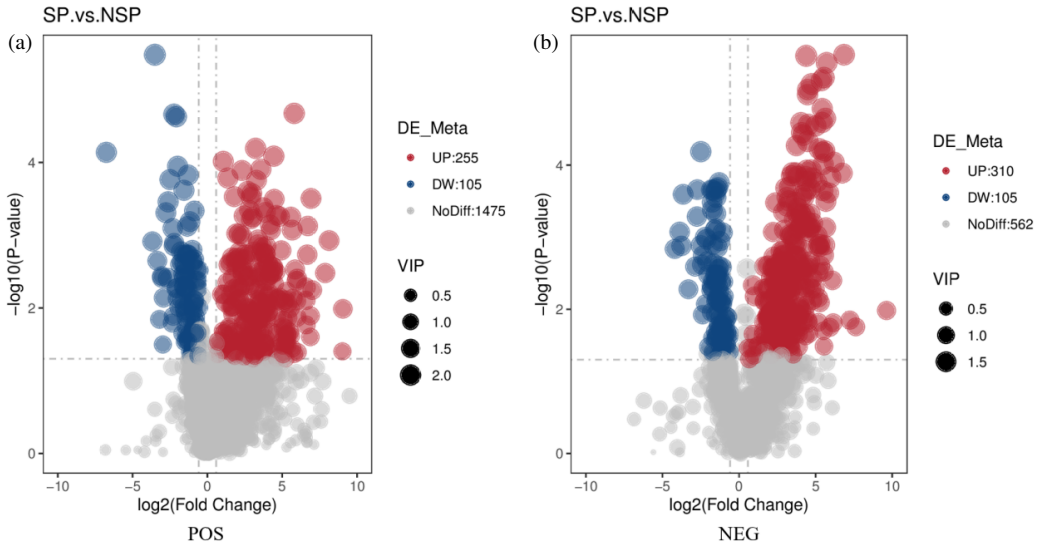


**FIGURE 7**. Volcano plot of differential metabolites in bronchoalveolar lavage fluid from patients with severe pneumonia vs. non-severe pneumonia controls.

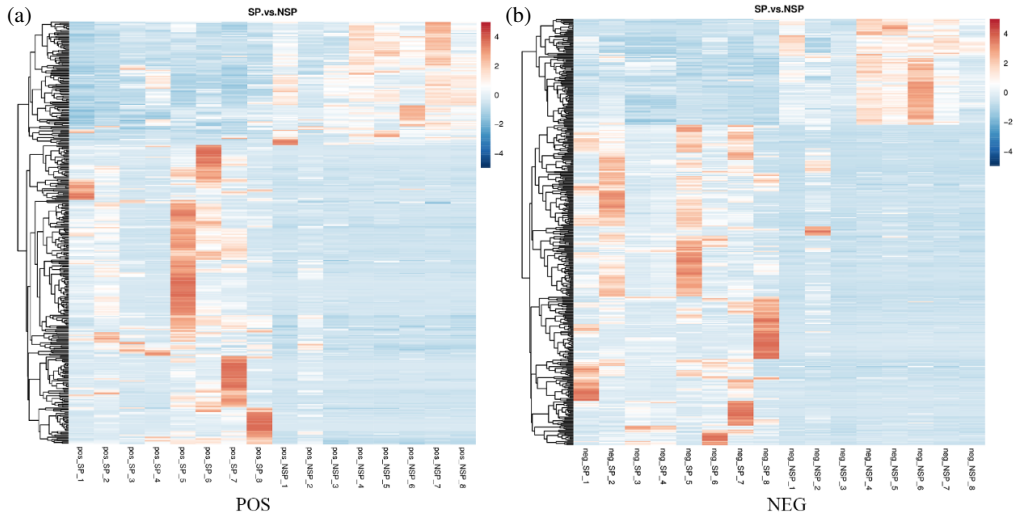


**FIGURE 8**. Clustering heatmap of differential metabolites in bronchoalveolar lavage fluid from patients with severe pneumonia vs. non-severe pneumonia controls.

This suggests that the lung metabolic state in severe pneumonia is not merely a simple worsening of non-severe pneumonia but is an independent pathological state with its own unique and stable biochemical fingerprint. This clear discriminatory ability provides initial evidence for these metabolites as reliable biomarkers for disease severity and also enables our later classification between severe pneumonia group and non-severe pneumonia group using fingerprint spectra of BALF.

An analysis and screening of differential metabolites was performed. As shown in the differential volcano plot (Figure 7), the severe pneumonia group exhibited upregulation of 255 positive ion metabolites and 310 negative ion metabolites, along with downregulation of 105 positive ion metabolites and 105 negative ion metabolites.

Subsequently, cluster analysis was performed on the differential metabolites. As shown in Figure 8, a horizontal comparison reveals significant differences in the clustering patterns of both positive and negative ion metabolites between the severe pneumonia and non-severe pneumonia groups.

KEGG pathway analysis was performed on the metabolites with differential expression. As shown in Figure 9, the pathways "Global and overview maps" and "Amino acid metabolism" exhibited the highest percentage of enriched positive and negative ion metabolites relative to the total number of metabolites. Specifically, in the "Global and overview maps" category, positive ion metabolites accounted for 26.79% and negative ion metabolites for 21.67%. In the "Amino acid metabolism" pathway, positive and negative ion metabolites represented 19.64% and 15%, respectively. Further analysis revealed that among positive ion metabolites, "Drug metabolism — cytochrome P450" and "ABC transporters" were the most enriched pathways, while among negative ion metabolites, "Pyruvate metabolism" and "Cysteine and

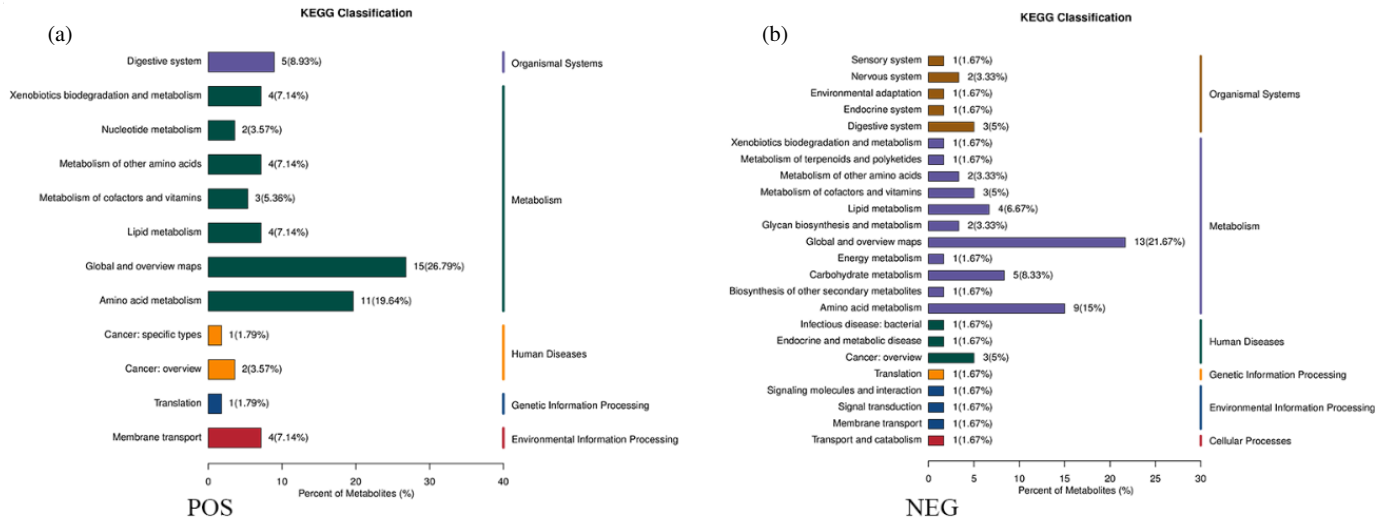


**FIGURE 9**. KEGG pathway analysis of differential metabolites in BALF from patients with severe pneumonia vs. non-severe pneumonia controls.

methionine metabolism" showed the highest enrichment. The enrichment of these specific pathways provides insights into the biochemical basis of severe pneumonia.

### 3.2.2. Classification of Severe Pneumonia Using Fingerprint Spectra of BALF

For the BALF samples, the Raman signal was very weak, and thus our home-made SERS chips were utilized to enhance the Raman signals. The SERS chip is a Metal-Insulator-Metal (MIM) structure, consisting of a top layer of Ag nanoparticles on an insulating layer ($SiO_2$, 50 nm, 0.02 nm/sec) backed with an Ag metal reflective mirror (Ag, 100 nm), see [11] for details. Figure 10 shows the SERS-enhanced Raman spectra of BALF from severe pneumonia (SP) patients versus non-severe pneumonia controls. The plot shows the difference in the average Raman signals between the two groups. The shaded areas represent the standard deviation ($\pm 1\sigma$).

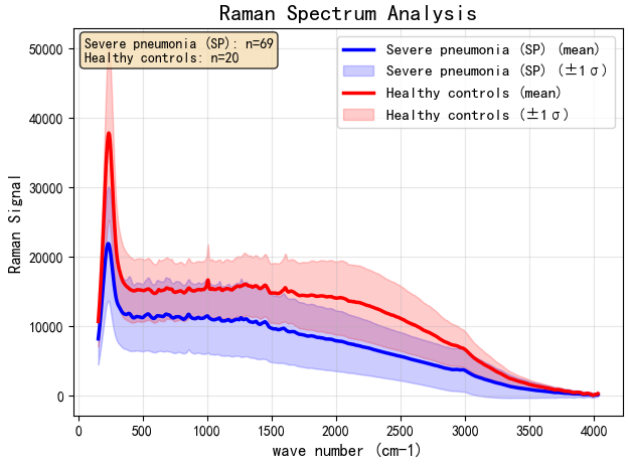


**FIGURE 10**. Raman spectra of bronchoalveolar lavage fluid (BALF) from severe pneumonia (SP) patients versus non-severe pneumonia controls.

The classification results for BALF Raman spectra are shown in Figure 11. The model achieved a training accuracy of $96.71\% \pm 1.86\%$ and a testing accuracy of $90.62\% \pm 3.95\%$.

Figure 12 shows the average infrared absorption spectral fingerprints of BALF samples from severe pneumonia patients and **non-severe pneumonia controls**. The shaded areas represent the standard deviation ($\pm 1\sigma$).

The confusion matrices of the classification results for BALF FTIR spectra are shown in Figure 13. The model achieved a training accuracy of $96.74\% \pm 4.05\%$ and a testing accuracy of $90.53\% \pm 7.74\%$.

## 4. DISCUSSIONS

This study demonstrates the significant clinical potential of spectroscopic fingerprinting for the rapid and accurate diagnosis of severe pneumonia. By focusing exclusively on spectral data and omitting traditional clinical markers, we have shown that these spectroscopic methods of molecular fingerprints can provide a highly accurate assessment of a patient's condition related to severe pneumonia. The analysis revealed several key differences in the spectral fingerprints of the two groups. Prominent peaks in the Raman spectra were observed around $\sim 1004\,\mathrm{cm}^{-1}$ (phenylalanine ring breathing), which reflects changes in protein composition due to inflammation, and at $\sim 1450\,\mathrm{cm}^{-1}$ ($-$CH2 deformation), which is indicative of alterations in lipid and protein content. Similarly, the FTIR spectra showed distinct differences in the Amide I band at $\sim 1650\,\mathrm{cm}^{-1}$ (related to protein secondary structure changes).

The most notable finding is the superior performance of FTIR spectroscopy for blood samples. The exceptional diagnostic precision (97.78% test accuracy) and stability (SD $<$ 0.02) make it a highly reliable tool for clinical applications. This high level of accuracy provides reliable technical basis for the development of rapid screening technologies. The ability to diagnose with such precision from a simple blood test (with a small amount of less than 0.05 mL) offers a compelling alternative to current time-consuming and less specific methods.

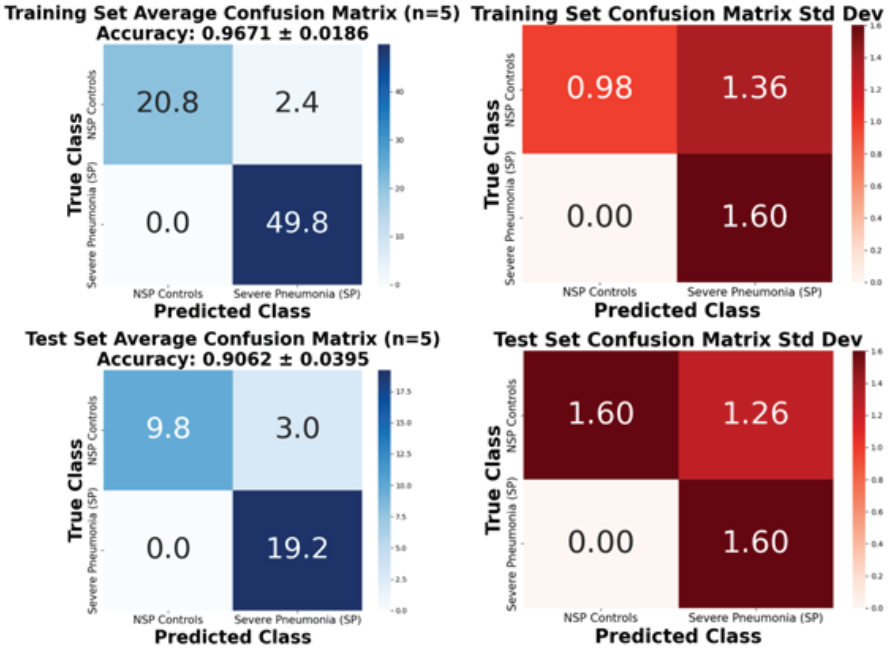


**FIGURE 11**. AI-based classification results using the Raman spectra of BALF.

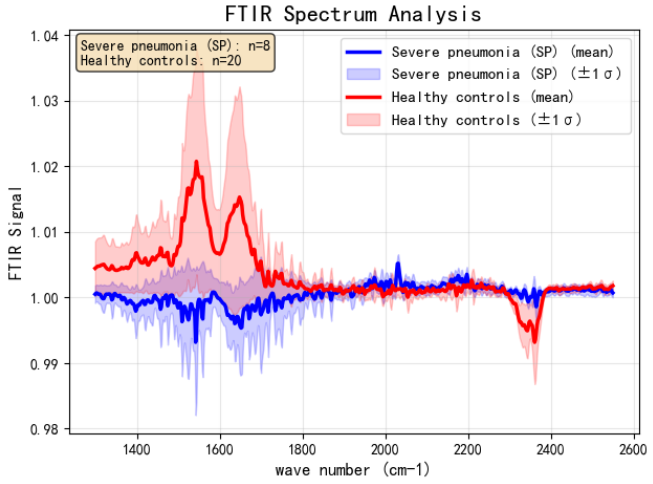


**FIGURE 12**. FTIR spectra of BALF from severe pneumonia (SP) patients versus non-severe pneumonia controls.

While Raman spectroscopy was less accurate and stable than FTIR for blood samples, its performance still exceeded 90% accuracy. Native blood signals were sufficient for accurate classification, eliminating the need for SERS enhancement. This simplifies the measurement process and reduces costs, making it a viable alternative, especially if a dedicated Raman system is already available. However, for biofluids with weak intrinsic signals like BALF, SERS enhancement is essential.

The use of blood samples represents a significant clinical advantage. Compared to BALF, which requires a more invasive procedure, blood is easily and routinely collected in emergency and critical care settings. This makes a blood-based spectral analysis method highly practical for clinical adoption.

**Continuous monitoring** is a key benefit of rapid spectral measurements. Unlike traditional blood biochemical tests, which provide a static test at a single point in time, spectroscopic measurements can be performed quickly and repeatedly to monitor dynamic changes in a patient's biochemical state. This capability is crucial for tracking the progression of severe pneumonia and for providing early warnings of deterioration, which is a major limitation of current assessment systems. The ability to continuously monitor deep-seated pathophysiological changes could be a paradigm shift in disease management. The label-free and rapid nature of these spectroscopic techniques makes them ideal for such continuous, real-time monitoring.

Our approach overcomes the major limitations of 'omics' research, which, while powerful, is not suitable for rapid clinical use due to its cost and time-intensive nature. The spectral fingerprinting technology demonstrated here provides a rapid, cost-effective, and highly accurate alternative that can bridge the gap between in-depth molecular research and timely clinical decision-making.

The same self-developed single-cell SERS-Raman imager [11] can also be used to detect the bacterial cells in BALF. The SERS-enhanced Raman spectra for different individual bacterial cells in the BALF of severe pneumonia are shown in Figure 14, where different types of bacterial cells have distinct Raman characteristics. Like our earlier work for single-cell Raman spectroscopy for rapid detection of bacteria in ballast water [11], together with an AI algorithm these Raman characteristics can be utilized to identify the types of bacterial cells.

We acknowledge that the present study, while demonstrating significant potential, has several limitations that will be addressed in future work. First, while the high accuracy of the classification models is promising, the study's generalizability is a key consideration. The research was conducted at a single clinical center in Taizhou Hospital with a specific patient cohort. This single-center design may limit the applicability of

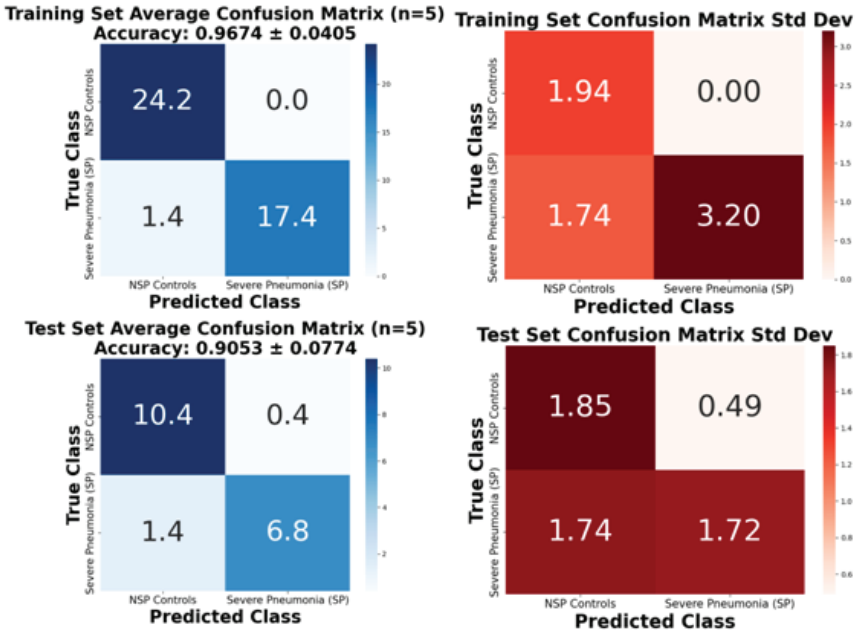


**FIGURE 13**. AI-based classification results using the FTIR spectra of BALF.

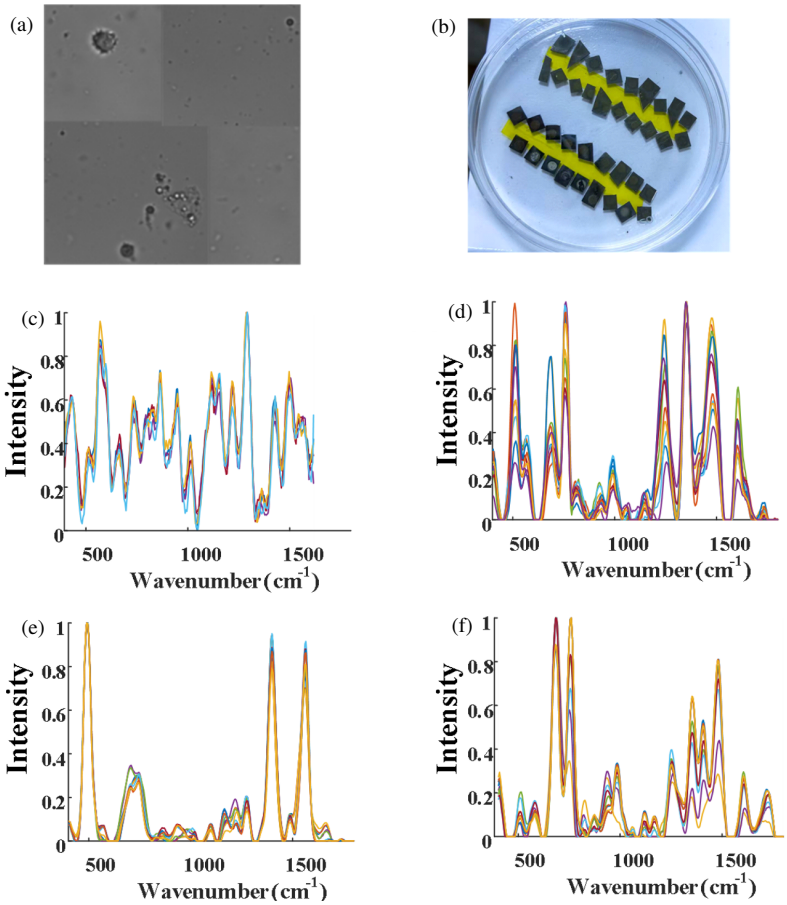


**FIGURE 14**. Raman spectra (detected with our self-developed confocal Raman imaging system with SERS) for different individual bacterial cells in BALF of severe pneumonia. (a) Microscopic image of BALF, where bacteria of Cryptococcus is marked with a blue box at the top-right corner; (b) Our SERS chips based on a metal-insulator-metal (MIM) structure; (c) Baseline-corrected SERS spectra of BALF samples; (d)–(f) Baseline-corrected normalized SERS spectra of individual bacterial cells of Cryptococcus, Acinetobacter baumannii, and Klebsiella pneumoniae.

the findings to a broader population, as patient demographics, disease presentation, and treatment protocols may vary across different geographic locations or healthcare systems. Second, a potential limitation lies in the data imbalance of both cohorts. The blood cohort, for instance, has a large disparity between the number of severe pneumonia patients (40) and healthy controls (138). While the computational model was specifically designed to mitigate overfitting and demonstrated consistent performance between training and testing sets, this inherent imbalance is a challenge that warrants a larger, more balanced cohort in future validation studies.

## 5. CONCLUSION

We successfully demonstrated the effectiveness of using blood spectral data with deep learning for the classification of severe pneumonia, without relying on any clinical blood biochemical test results. This approach is a critical departure from our prior work, which combined spectral data with external clinical biochemical test results. The present study's exclusive focus on spectral data is a crucial step toward developing a diagnostic tool that can be used for continuous, real-time monitoring in a critical care setting, a capability that is not feasible with traditional clinical tests that provide only a static snapshot. Our comprehensive evaluation of two spectroscopic modalities, Raman and FTIR, across different biofluids clearly indicates that while both methods are highly effective, their performance varies significantly depending on the biofluid being analyzed. FTIR spectroscopy demonstrated higher accuracy for blood samples (97.78%), while Raman, particularly when enhanced with SERS, was more effective for BALF (90.62%). This explains why both techniques were employed — to provide a comprehensive and reliable diagnostic solution across different sample types. The deep learning approach using blood spectral data shows significant clinical potential for pneumonia diagnosis. This high accuracy, coupled with the rapid and label-free nature of blood-based testing, provides a reliable technical foundation for developing future diagnostic tools. The comprehensive head-to-head comparison of two different spectroscopic modalities (Raman and FTIR) on two distinct biofluids (blood and BALF) provides valuable and novel insights into the practical applicability and performance trade-offs of these techniques in a real-world clinical context. This comparative analysis provides a reliable technical foundation for developing future diagnostic tools. The single-cell SERS-Raman imager can be used to identify the types of infecting bacterial cells in BALF as well as the severity of pneumonia through a single measurement.

This approach has the potential to overcome the limitations of current diagnostic methods, shorten the time to diagnosis, and enable continuous monitoring of a patient's condition, ultimately improving outcomes for patients with severe pneumonia.

Note that machine-learning-enabled rapid diagnostics is in fast progress (see, e.g., [12–14]), and AI-classification with multiple measurement data obtained with different types of equipment may give better accuracy.

## ACKNOWLEDGEMENT

This work was jointly supported by "Pioneer" and "Leading Goose" R&D Program of Zhejiang Province (2022C03051), the Ningbo Science and Technology Project (grant No. 2023Z122), The National Natural Science Foundation of China (W2412107), and the Joint Fund of Zhejiang Provincial Natural Science Foundation of China under Grant No. LKLY25H200001. The authors gratefully thank Yongbo Jiang and Yufeng Zheng of Taizhou Hospital for providing some samples.

## REFERENCES

[1] Chernii, V. I., "Severe community-acquired pneumonia: Principles of diagnostics and intensive therapy," *Infusion & Chemotherapy*, Vol. 6, No. 3, 7–15, 2023.

[2] Niederman, M. S. and A. Torres, "Severe community-acquired pneumonia," *European Respiratory Review*, Vol. 31, No. 166, 220123, 2022.

[3] Woodhead, M., C. A. Welch, D. A. Harrison, G. Bellingan, and J. G. Ayres, "Community-acquired pneumonia on the intensive care unit: Secondary analysis of 17,869 cases in the ICNARC Case Mix Programme Database," *Critical Care*, Vol. 10, No. Suppl 2, S1, 2006.

[4] Huang, T., X. Zheng, L. He, and Z. Chen, "Diagnostic value of deep learning-based CT feature for severe pulmonary infection," *Journal of Healthcare Engineering*, Vol. 2021, No. 1, 5359084, Nov. 2021.

[5] Cui, K., D. Gong, X. Chen, Y. Xu, H. Li, Y. Zhu, J. Evans, X. Gong, Z. Shi, Y. Xu, and S. He, "Classification of severe bacterial pneumonia based on CT images and deep learning," *Progress In Electromagnetics Research Letters*, Vol. 127, 23–28, 2025.

[6] Martin, C., K. S. Mahan, T. D. Wiggen, A. J. Gilbertsen, M. I. Hertz, R. C. Hunter, and R. A. Quinn, "Microbiome and metabolome patterns after lung transplantation reflect underlying disease and chronic lung allograft dysfunction," *Microbiome*, Vol. 12, No. 1, 196, 2024.

[7] Cao, X. E., S. Y. Ongagna-Yhombi, R. Wang, Y. Ren, B. Srinivasan, J. A. Hayden, Z. Zhao, D. Erickson, and S. Mehta, "A diagnostic platform for rapid, simultaneous quantification of procalcitonin and C-reactive protein in human serum," *Ebiomedicine*, Vol. 76, 103867, 2022.

[8] Farooq, S., M. Del-Valle, M. O. D. Santos, S. N. D. Santos, E. S. Bernardes, and D. M. Zezell, "Rapid identification of breast cancer subtypes using micro-FTIR and machine learning methods," *Applied Optics*, Vol. 62, No. 8, C80–C87, 2023.

[9] Zhang, C., J. Li, W. Luo, and S. He, "Ai-assisted detection for early screening of acute myeloid leukemia using infrared spectra and clinical biochemical reports of blood," *Bioengineering*, Vol. 12, No. 4, 340, 2025.

[10] Jiao, C., J. Liao, and S. He, "An aberration-free line scan confocal Raman imager and type classification and distribution detection of microplastics," *Journal of Hazardous Materials*, Vol. 470, 134191, 2024.

[11] Yang, A., Z. Hu, X. Zou, Y. Zhang, J. Qian, S. Li, J. Liang, and S. He, "Single-cell Raman spectroscopy for rapid detection of bacteria in ballast water and UV254 treatment evaluation," *Talanta*, Vol. 284, 127266, 2025.

[12] Han, G.-R., A. Goncharov, M. Eryilmaz, S. Ye, B. Palanisamy, R. Ghosh, F. Lisi, E. Rogers, D. Guzman, D. Yigci, *et al.*, "Machine learning in point-of-care testing: Innovations, challenges,

and opportunities," *Nature Communications*, Vol. 16, No. 1, 3165, 2025.

[13] Amin, Y., P. Cecere, T. Pomili, and P. P. Pompa, "Smartphone-integrated YOLOv4-CNN approach for rapid and accurate point-of-care colorimetric antioxidant testing in saliva," *Progress In Electromagnetics Research*, Vol. 181, 9–19, 2024.

[14] Zhu, H., J. Luo, J. Liao, and S. He, "High-accuracy rapid identification and classification of mixed bacteria using hyperspectral transmission microscopic imaging and machine learning," *Progress In Electromagnetics Research*, Vol. 178, 49–62, 2023.