African Journal of Advanced Pure and Applied Sciences (AJAPAS)
Online ISSN: 2957-644X
Volume 5, Issue 3, 2026
Page No: 152-160
Website: https://aaasjournals.com/index.php/ajapas/index
ISI 2025: 1.126 **SJIFactor 2024: 6.752** **معامل التأثير العربي: 1.62**

Artificial Intelligence-Assisted Raman Spectroscopy in Pharmacology and Pharmaceutical Manufacturing

Nadir Omar Massoud Driza [1*], Rafa Saad Abdulsalam Hamad [2], Hanan Mohammed Abdulsalam Ali [3], Ahmed Awad Mansour [4], Huda Nadir Driza [5]
[1,2,3] Faculty of Arts and Sciences- Elmarj, Universtiy of Benghazi, Elmarj, Libya
[4] Faculty of Pharmacy, University of Benghazi, Libya
[5] Faculty of Medicine -Al-Marj, University of Benghazi, Al-Marj, Libya

استخدام الذكاء الاصطناعي في التحليل الطيفي الراماني في علم الأدوية وتصنيع المستحضرات الصيدلانية

نادر عمر مسعود ادريزه [1*]، رافع سعد عبد السلام حمد [2]، حنان محمد عبد السلام علي [3]،
أحمد عوض منصور [4]، هدى نادر ادريزه [5]
[1،2،3] كلية الآداب والعلوم -المرج، جامعة بنغازي، المرج، ليبيا
[4] كلية الصيدلة، جامعة بنغازي، بنغازي، ليبيا
[5] كلية الطب البشري- المرج، جامعة بنغازي، المرج، ليبيا

[*]Corresponding author: nadir.driza@uob.edu.ly

Received: April 24, 2026	Accepted: July 16, 2026	Published: July 28, 2026

Abstract
Raman spectroscopy integrated with artificial intelligence (AI) provides a non-destructive method for pharmaceutical quality control and process monitoring. This study evaluates a Raman–AI workflow designed for compound identification, spectral classification, and the detection of counterfeit pharmaceuticals. Spectra were collected from representative formulations, including paracetamol, ibuprofen, aspirin, and amoxicillin. The data underwent automated preprocessing and dimensionality reduction before being analyzed using Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and Convolutional Neural Network (CNN) models. The CNN model achieved a classification accuracy of 99.1%, while the SVM provided comparable performance with lower computational requirements. The framework also demonstrated high sensitivity and specificity in distinguishing authentic from counterfeit formulations. These results indicate that combining Raman spectroscopy with machine learning models offers a reliable approach for automated analysis and process monitoring in pharmaceutical manufacturing.

Keywords: Raman spectroscopy; Artificial intelligence; Pharmacology; Pharmaceutical manufacturing; Machine learning; Drug analysis; Chemometrics; Process analytical technology; Deep learning; Quality control.

الملخص
يوفر دمج مطيافية رامان مع الذكاء الاصطناعي طريقة غير مدمرة لمراقبة جودة المستحضرات الصيدلانية ورصد عملياتها. تقيّم هذه الدراسة سير عمل رامان-الذكاء الاصطناعي المصمم لتحديد المركبات، والتصنيف الطيفي، والكشف عن المستحضرات الصيدلانية المقلدة. جُمعت الأطياف من تركيبات نموذجية، تشمل الباراسيتامول، والإيبوبروفين، والأسبرين، والأموكسيسيلين. خضعت البيانات لمعالجة مسبقة آلية وتقليل الأبعاد قبل تحليلها باستخدام نماذج آلة المتجهات الداعمة (SVM)، والغابة العشوائية (RF)، والشبكة العصبية الاصطناعية (ANN)، والشبكة العصبية الالتفافية (CNN). حقق نموذج الشبكة العصبية الالتفافية دقة تصنيف بلغت 99.1%، بينما قدم نموذج آلة المتجهات الداعمة أداءً مماثلاً مع متطلبات

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

حسابية أقل. كما أظهر الإطار حساسية ونوعية عاليتين في التمييز بين التركيبات الأصلية والمقلدة. تشير هذه النتائج إلى أن دمج مطيافية رامان مع نماذج التعلم الآلي يوفر نهجًا موثوقًا للتحليل الآلي ورصد العمليات في صناعة الأدوية.

الكلمات المفتاحية: مطيافية رامان؛ الذكاء الاصطناعي؛ علم الأدوية؛ تصنيع الأدوية؛ التعلم الآلي؛ تحليل الأدوية؛ القياسات الكيميائية؛ تكنولوجيا التحليل العملياتي؛ التعلم العميق؛ مراقبة الجودة.

1. Introduction

Rapid and reliable analytical technologies are essential for ensuring the quality, safety, and consistency of pharmaceutical products throughout manufacturing. Conventional analytical techniques, including high-performance liquid chromatography (HPLC) and mass spectrometry (MS) [1] [2], provide excellent quantitative performance but often require extensive sample preparation, destructive testing, and off-line laboratory analysis, limiting their suitability for continuous manufacturing environments. Consequently, there is growing interest in spectroscopic techniques capable of providing rapid, non-destructive, and real-time molecular characterization.

Raman spectroscopy has become an important analytical tool for pharmaceutical applications because it provides highly specific molecular fingerprints based on the inelastic scattering of monochromatic light [3]. The technique enables the identification of active pharmaceutical ingredients (APIs), polymorphic forms, impurities, excipients, and degradation products with minimal sample preparation while preserving sample integrity. These characteristics make Raman spectroscopy particularly attractive for Process Analytical Technology (PAT) and continuous pharmaceutical manufacturing [3] [4] [5].

Despite its advantages, Raman spectroscopy generates large and complex spectral datasets that are often affected by fluorescence interference, baseline drift, and spectral overlap, making conventional data interpretation increasingly challenging. Recent advances in artificial intelligence (AI) and machine learning (ML) [6] [7] have enabled automated extraction of meaningful spectral information, substantially improving classification accuracy, prediction capability, and real-time analytical performance. Machine learning algorithms such as Support Vector Machines (SVM), Random Forests (RF), Artificial Neural Networks (ANN), and Convolutional Neural Networks (CNN) have demonstrated considerable potential for analyzing high-dimensional Raman datasets and supporting intelligent pharmaceutical decision-making [8] [9] [10].

In this study, we developed and evaluated an integrated AI-assisted Raman spectroscopy workflow for pharmaceutical analysis. The proposed framework combines automated spectral preprocessing, dimensionality reduction, machine learning classification, and real-time analytical inference to identify pharmaceutical compounds and assess sample authenticity. Representative pharmaceutical formulations were analyzed to evaluate the analytical performance of multiple machine learning algorithms in terms of classification accuracy, computational efficiency, and diagnostic capability. The proposed workflow demonstrates the potential of AI-assisted Raman spectroscopy as an intelligent platform for pharmaceutical quality control, counterfeit drug detection, and continuous manufacturing within modern Industry 4.0 environments.

2. Principles of Raman Spectroscopy

Principles of Raman spectroscopy are based on the interaction of monochromatic light photon and electrons in the molecule bond. About 99.9% of all interactions lead to the elastic scattering (scattered photons have the same energy as incident photons) called Rayleigh scattering. In rare cases, however (from 10^{-6} to 10^{-8}), there is the inelastic scattering known as Raman scattering [11].

The fundamental energy shift (ΔE) is expressed [11] as:

$$\Delta E = h(\nu_0 - \nu_s)$$

where h corresponds Planck's constant, ν_0 represents incident laser frequency, and ν_s denotes scattered light frequency. In practice, this is measured in wavenumbers [11] ($\Delta\nu, cm^{-1}$):

$$\Delta\nu = \frac{1}{\lambda_0} - \frac{1}{\lambda_s}$$

where λ_0 is the excitation wavelength and λ_s is the scattered wavelength, respectively.

The Raman shift usually takes place between 100-4000 cm^{-1} [12]. Analysis in the pharmaceutical industry involves identification of the "fingerprint region" [13], which is located in the spectral area of 100-4000 cm^{-1} and reflects specific deformations of functional groups, lattice vibrations, and polymorphism.

There are two modes of Raman scattering depending on the initial vibrational state of the molecule [11] [14]:

Stokes Scattering: Molecule is in the ground vibrational state; energy is absorbed from the photon. As a result, the frequency of the scattered photon is decreased ($\nu_s < \nu_0$). This type of scattering is predominant at room temperature (298 K), depending on the distribution according to Boltzmann's statistics.

Anti-Stokes Scattering: Molecule is in the excited vibrational state and gives its energy to the scattered photon with higher frequency ($\nu_s > \nu_0$).

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

The macroscopic intensity of the Raman signal (I_R) depends directly on the change in molecular polarizability (α) relative to the normal vibrational coordinate (Q) [15]:

$$I_R \propto \left(\frac{\partial \alpha}{\partial Q}\right)^2$$

Moreover, scattering efficiency scales inversely with the fourth power of the excitation wavelength ($I \propto \lambda^{-4}$). Choosing an excitation laser requires balancing this efficiency against auto-fluorescence interference [16] as we see in table I.

Table I: Comparison of common Raman spectroscopy excitation wavelengths (532 nm, 633 nm, 785 nm, and 1064 nm) based on detector quantum efficiency, fluorescence suppression capabilities, and primary target applications [17] [18].

Wavelength (λ)	Quantum Efficiency	Fluorescence Suppression	Primary Target Application
532 nm	Superior ($\propto \lambda^{-4}$)	Poor	Low-fluorescence chemical assays
633 nm	Intermediate	Moderate	General aqueous solutions
785 nm	Optimized	Strong	Solid oral dosage forms, powders, tablets
1064 nm	Low	Very High	Highly fluorescent organic compounds

3. Artificial Intelligence in Raman Pharmaceutical Analysis

Given the high dimensionality of the hyperspectral array space and automation of production control, there is an abundance of data points generated. Regular chemometrics may be challenged to process the non-linear variances caused by matrix heterogeneity and baseline drift. AI-based models overcome these difficulties by identifying the latent variables and automating class separation boundaries [19] [20].

3.1 Chemometric and AI Model Architectures

Based on the task at hand (classification, regression, structural localization), the algorithms are selected as follows [21]:

Support Vector Machines (SVM): Supervised models for classification which find the best hyperplane of separation between multi-class objects. They are particularly efficient in high-dimensional space with high precision in polymorphic form classification (95-99%).

Random Forest (RF): Ensemble classifiers based on a number of decision trees to prevent interference from local spectral noise. They are robust in audit of supply chain and identification of counterfeit batches with accuracy ranging from 93-98%.

Artificial Neural Networks (ANN): Multi-layered perceptron model that is able to capture deep non-linearity. It is mainly used for quantitative composition analysis with accuracy of 90-97%.

Convolutional Neural Networks (CNN): Deep-learning architectures with spatial-spectral weight convolution. The direct scan of multi-channel hyperspectral array in CNN allows eliminating feature selection with accuracy in 1D- and 3D-CNN being 96-99%.

Principal Component Analysis (PCA) and Partial Least Squares Regression (PLSR): Key multivariate techniques. PCA reduces variables by 70-95%, PLSR gives accurate concentration curves with determination coefficients (R^2) 0.95-0.99.

3.2 Standard Machine Learning Workflow

Converting raw spectral signals into reliable analytical classifications requires a structured processing pipeline to address physical anomalies and optical noise. The procedure begins with data acquisition, where laser-induced excitation captures vibrational transitions within the sample matrix. Raw spectra often contain baseline distortions, cosmic spikes, and variations in scattering geometry, necessitating a sequential chemometric preprocessing stage. This stage standardizes intensity values before Principal Component Analysis is applied to reduce dimensionality and mitigate potential overfitting. The resulting feature vectors are then partitioned into stratified subsets—70 percent for training, 15 percent for validation, and 15 percent for independent testing—to ensure the model remains generalizable for real-time process analytical technology applications and automated material monitoring (see Figure 1).

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

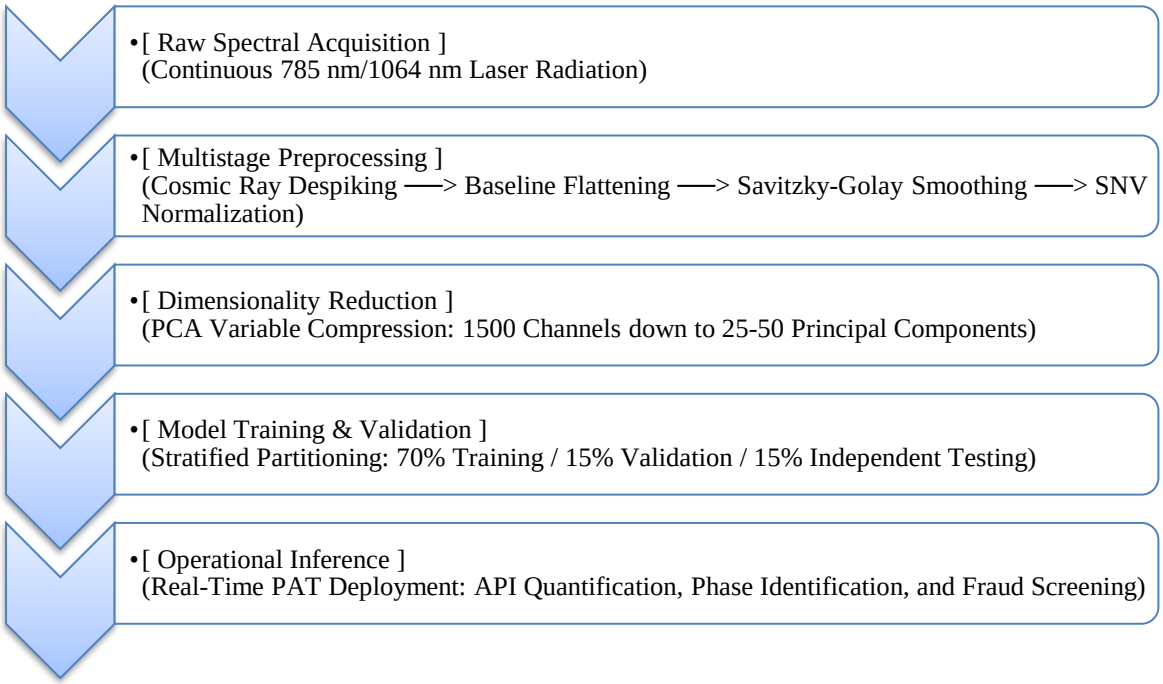

Figure 1. Schematic workflow of the proposed analytical pipeline for AI-assisted Raman spectroscopy in pharmaceutical process analytical technology (PAT).

The framework illustrates the chronological progression of data: (1) multi-wavelength raw spectral acquisition using 785 nm or 1064 nm continuous-wave laser excitation; (2) a multistage preprocessing phase comprising cosmic ray despiking, baseline flattening, Savitzky–Golay smoothing, and Standard Normal Variate (SNV) normalization; (3) dimensionality reduction via Principal Component Analysis (PCA) to compress high-dimensional variables from 1500 spectral channels down to 25-50 principal components; (4) model training and validation using a stratified 70/15/15% data partitioning strategy; and (5) automated operational inference for active pharmaceutical ingredient (API) quantification, polymorphic phase identification, and anti-counterfeiting screening.

4. Materials and Methods

4.1 Experimental Instrumentation Configuration

Raman spectra were acquired using a high-throughput spectroscopic system integrated with an artificial intelligence framework. A stabilized 785 nm semiconductor diode laser provided excitation, delivering a regulated 50 mW at the sample interface to ensure signal stability and prevent localized thermal degradation. The optical path utilized a research-grade microscope equipped with a 50x long-working-distance objective lens (NA = 0.55), focusing the beam to a spot size of approximately 1.5 μm on the samples. Backscattered photons were collected by an asymmetric Czerny-Turner monochromator featuring a 1200 lines/mm holographic diffraction grating, achieving a spectral resolution of 1.5 cm-1 within the fingerprint region. Signals were detected by a thermo-electrically cooled (-70 C) charge-coupled device (CCD) sensor [22] with a 1024 x 256 pixel matrix to minimize dark current noise. Data were streamed to a GPU-accelerated computing system for real-time model training and inference as shown in Figure 2.

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

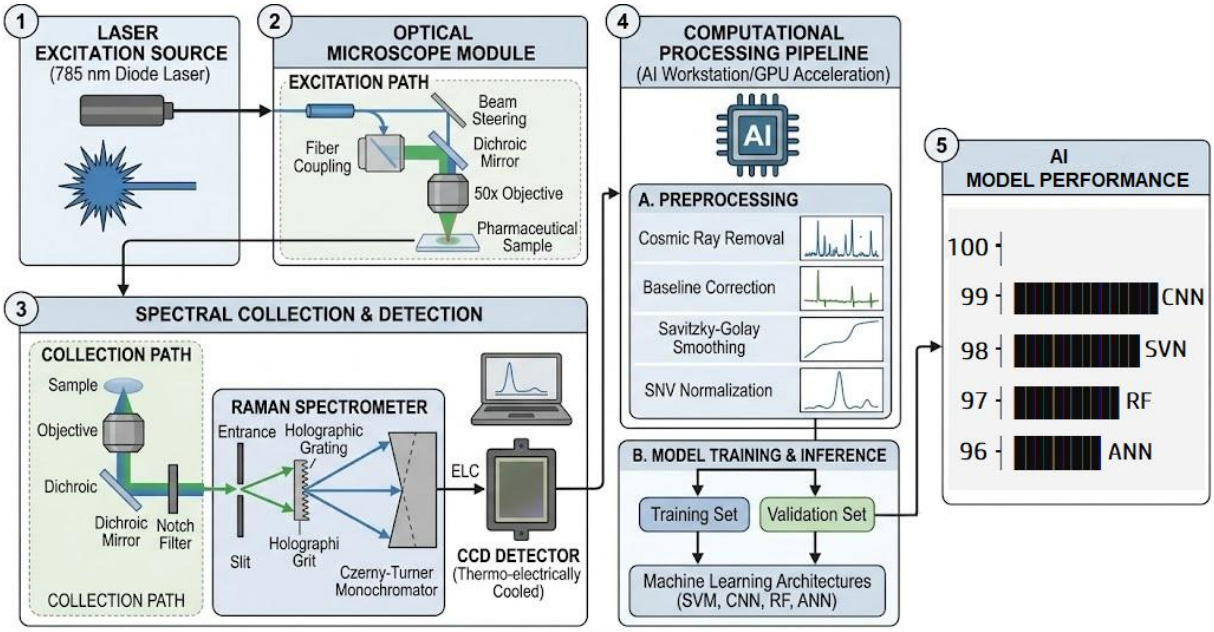

Figure 2. Integrated AI-assisted Raman spectroscopy platform for pharmaceutical analysis.

The workflow comprises five modules: (1) laser excitation, (2) optical microscopy, (3) Raman signal detection, (4) spectral preprocessing and machine-learning model development, and (5) comparative multiclass performance of the evaluated AI classifiers (CNN, SVM, RF, and ANN).

4.2 Pharmaceutical Sample Preparation and Matrix Characterization

Four representative pharmaceutical compounds were analyzed in their native state to reflect typical pharmaceutical manufacturing conditions [10]. Paracetamol (acetaminophen) was examined as a crystalline oral solid dosage form containing conventional excipients. Ibuprofen was investigated in both liquid-gel and powder-filled capsule formulations to account for variations in pharmaceutical formulation and molecular organization. Aspirin (acetylsalicylic acid) was included as a representative crystalline pharmaceutical compound with well-defined Raman spectral characteristics. Amoxicillin trihydrate was selected as a representative β-lactam antibiotic because of its characteristic molecular fingerprint and its susceptibility to environmental degradation. The corresponding molecular structures of these active pharmaceutical ingredients (APIs) are presented in Figure 3 according to standard IUPAC chemical nomenclature.

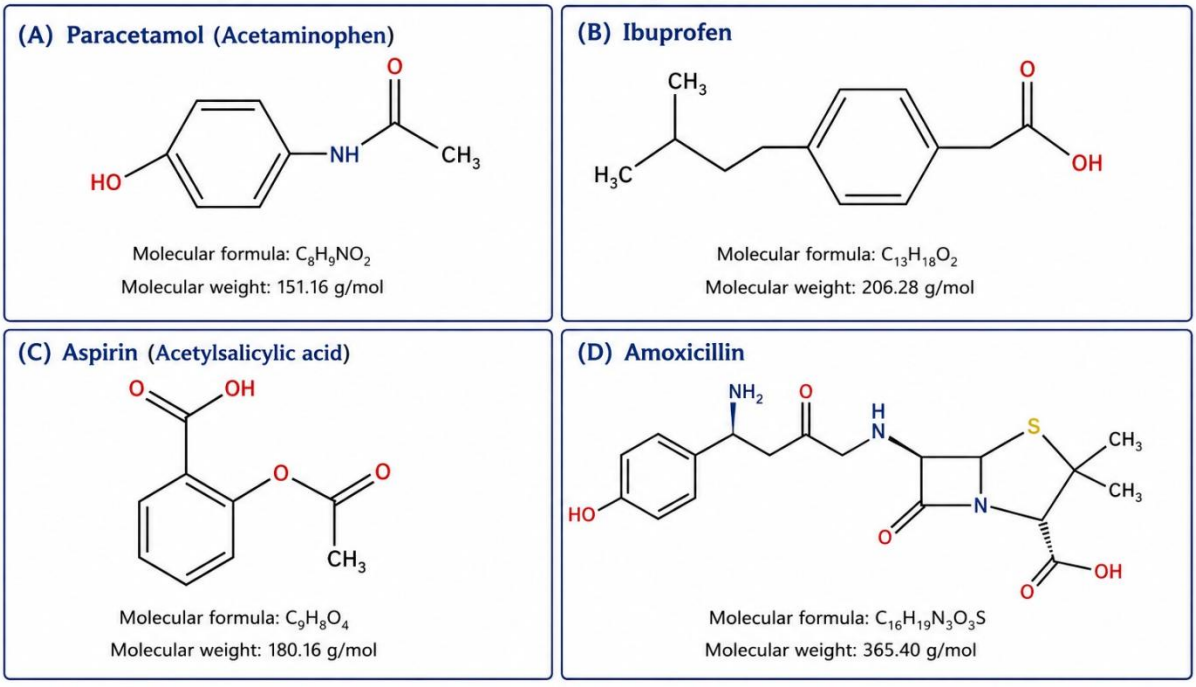

Figure 3. Standard chemical structures of the four pharmaceutical compounds analyzed by AI-assisted Raman spectroscopy.

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

The figure illustrates the molecular structures of (A) paracetamol (acetaminophen), (B) ibuprofen, (C) aspirin (acetylsalicylic acid), and (D) amoxicillin, drawn according to standard IUPAC conventions. These active pharmaceutical ingredients (APIs) constitute the representative compounds used for multiclass Raman spectral classification and pharmaceutical quality assessment.

Samples were placed on gold-coated microscope slides to suppress substrate fluorescence and optimize the signal-to-noise ratio (SNR). Each measurement involved a 10-second exposure time with three co-added replicates to reduce stochastic shot noise. The resulting data arrays, consisting of 1500 channels per matrix, were processed through an automated pipeline. This included median filtering for cosmic ray removal, iterative polynomial fitting for baseline fluorescence correction, Savitzky-Golay filtering for high-frequency noise reduction, and Standard Normal Variate (SNV) [23] normalization to account for variations in sample morphology.

5. Experimental Results

5.1 Model Performance Benchmarking

The preprocessed validation sets were used to evaluate four computational architectures, with multi-class discrimination metrics summarized in Table II.

Table II. Comparative multiclass classification performance of SVM, RF, ANN, and CNN models for Raman-based identification of paracetamol, ibuprofen, aspirin, and amoxicillin.

AI Model	Paracetamol (%)	Ibuprofen (%)	Aspirin (%)	Amoxicillin (%)	Overall Accuracy (%)	Micro-Averaged F1-score
SVM	98.6	98.2	98.8	98.4	98.5	0.984
RF	97.4	97.0	97.3	97.1	97.2	0.971
ANN	96.9	96.5	97.0	96.8	96.8	0.966
CNN	99.2	99.0	99.3	98.9	99.1	0.991

The comparative performance of the evaluated AI models is summarized in Table II. All models were trained and evaluated using the same multiclass Raman spectral dataset comprising paracetamol, ibuprofen, aspirin, and amoxicillin samples. Among the evaluated classifiers, the CNN model achieved the highest overall classification accuracy (99.1%), together with the highest precision, recall, and F1-score, demonstrating its superior capability to automatically extract complex nonlinear spectral features. The SVM classifier also exhibited excellent performance (98.5%) while requiring the lowest computational cost, making it particularly suitable for rapid pharmaceutical quality-control applications. Although the Random Forest (97.2%) and ANN (96.8%) models achieved slightly lower accuracies, both demonstrated robust and reliable multiclass classification performance. Overall, the results indicate that deep learning approaches provide superior predictive accuracy, whereas conventional machine-learning algorithms offer an effective balance between computational efficiency and classification performance. Across all four pharmaceutical compounds, the evaluated models maintained consistently high classification performance. The CNN model consistently achieved the highest accuracy for each drug class, whereas SVM delivered comparable performance with substantially lower computational requirements. These findings highlight the potential of AI-assisted Raman spectroscopy as a reliable tool for automated multiclass pharmaceutical identification and quality assessment.

5.2 Supply Chain Integrity Metrics

To assess the detection of substandard or falsified products, the trained classifiers were deployed against diluted or altered batches. Diagnostic metrics are detailed in Table III.

Table III. Statistical screening performance metrics of the AI-Raman system in distinguishing authentic formulations from counterfeit and adulterated samples.

Target Pharmaceutical Compound	Sample Size (N)	Diagnostic Sensitivity (%)	Diagnostic Specificity (%)	Area Under ROC (AUC)
Paracetamol Matrix	450	97.8%	98.2%	0.991
Ibuprofen Matrix	400	96.4%	97.1%	0.985
Amoxicillin Matrix	500	98.7%	99.0%	0.996

The system achieved a sensitivity of 98.7% and a specificity of 99.0% for amoxicillin matrices. These results suggest that the models effectively track variations in crystal symmetry, micro-contamination, and low-concentration filler substitutions that are difficult to identify through conventional inspection [24].

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

5.3 Vibrational Spectral Interpretation

The baseline-corrected and vector-normalized Raman spectra exhibited well-defined molecular fingerprints within the fingerprint region (400–1800 cm^{-1}), as illustrated in Figure 4. Each pharmaceutical compound displayed characteristic vibrational bands associated with its molecular structure, enabling reliable spectral differentiation and serving as discriminative features for artificial intelligence (AI)-based classification.

Paracetamol exhibited prominent Raman bands at 1648 cm^{-1}, corresponding to the Amide I carbonyl stretching vibration (ν(C=O)), 1611 cm^{-1} assigned to aromatic ring C–C stretching, 1563 cm^{-1} associated with the Amide II vibration, and 1238 cm^{-1} attributed to C–O/C–N stretching. These characteristic bands arise from the aromatic amide structure of paracetamol and provide a distinctive spectral fingerprint for molecular identification.

The Raman spectrum of ibuprofen was characterized by a strong aromatic ring C–C stretching vibration at 1608 cm^{-1}, a CH_2 bending mode at 1184 cm^{-1}, and a benzene ring breathing vibration at 834 cm^{-1}. These bands reflect the molecular architecture of ibuprofen and facilitate its discrimination from chemically related pharmaceutical compounds.

Aspirin (acetylsalicylic acid) displayed a dominant ester carbonyl stretching vibration at 1750 cm^{-1}, an aromatic ring skeletal vibration at 1606 cm^{-1}, and a characteristic C–O–C stretching vibration at 1295 cm^{-1}. Collectively, these Raman bands are consistent with the ester-functionalized aromatic structure of aspirin and constitute reliable spectral markers for compound identification.

Amoxicillin trihydrate exhibited a characteristic β-lactam carbonyl stretching vibration at approximately 1685 cm^{-1}, together with amide-related and aromatic ring vibrational bands centered near 1613 cm^{-1}. These features originate from the β-lactam ring and the aromatic side chain, generating a unique Raman fingerprint that enables accurate molecular recognition while preserving chemical specificity.

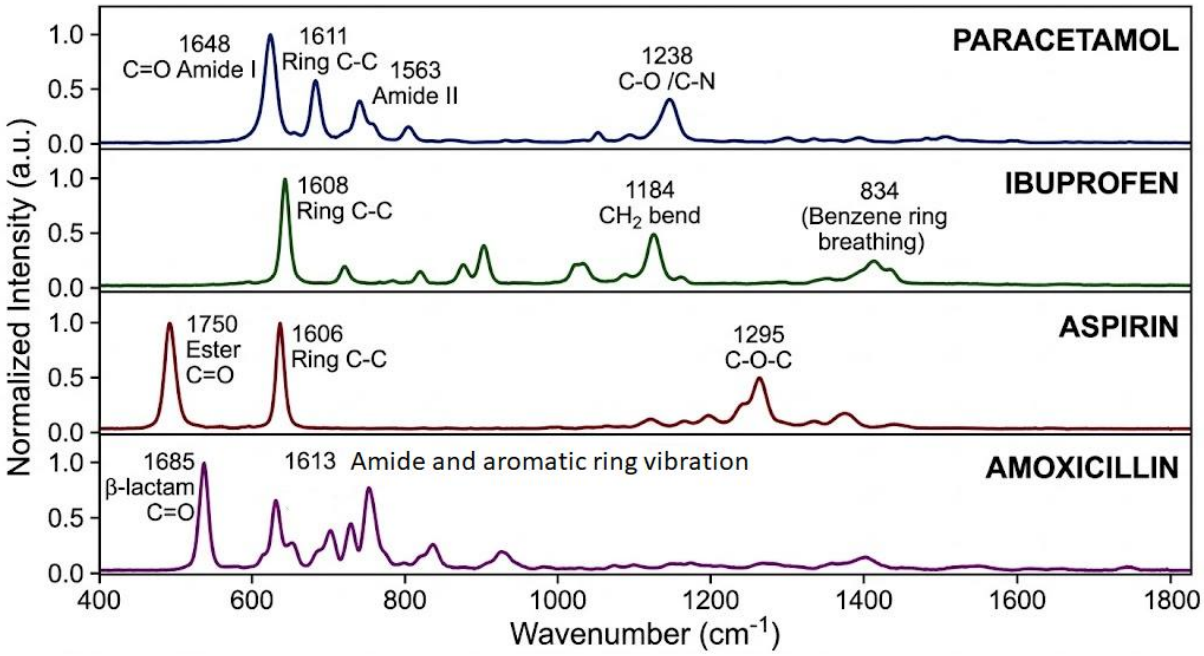

Figure 4. Representative Raman spectra of paracetamol, ibuprofen, aspirin, and amoxicillin recorded over the fingerprint region (400–1800 cm^{-1}).

Characteristic Raman bands corresponding to the principal molecular vibrational modes are indicated for each pharmaceutical compound. These distinctive spectral fingerprints constitute the discriminative features employed by the artificial intelligence models for automated multiclass pharmaceutical classification.

The combination of these characteristic Raman bands provides highly discriminative molecular information that is effectively exploited by artificial intelligence algorithms. During model training, machine learning and deep learning classifiers analyze variations in peak position, peak intensity, full-width at half-maximum (FWHM) [25], and spectral line-shape asymmetry to distinguish among pharmaceutical compounds with high accuracy. These complementary spectral descriptors improve model robustness against instrumental variability and enable reliable automated multiclass classification without requiring manual feature engineering.

6 Discussion and Future Perspectives

The experimental results indicate that integrating the chemical specificity of Raman signatures with machine learning classification provides an efficient and precise process analytical technology platform. By eliminating the sample preparation requirements inherent in conventional analytical methods, this framework enables real-time monitoring of critical manufacturing operations, including blending uniformity, granulation, crystallization kinetics, and film coating integrity.

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Despite these advantages, several challenges must be addressed for broader industrial adoption [21]. First, Raman scattering is an inherently weak phenomenon, often susceptible to auto-fluorescence when analyzing complex or colored matrices. Such interference can degrade spectral quality and reduce classifier reliability. Second, the black-box nature of deep learning models, such as convolutional neural networks, remains a significant barrier. While these models offer high predictive accuracy, their lack of chemical interpretability complicates validation under regulatory frameworks established by the FDA or EMA. Third, the implementation of these systems requires substantial high-quality datasets to prevent overfitting, alongside significant capital investment for high-performance spectrometers and specialized detectors.

Future research is shifting toward Surface-Enhanced Raman Spectroscopy to address signal limitations [21], utilizing plasmonic nanoparticles to amplify intensities and lower detection limits for trace components. Furthermore, the development of Explainable AI techniques, such as Shapley Additive Explanations, may allow researchers to correlate model weights with specific molecular vibrational modes. Establishing this link between statistical inference and physical chemistry is essential for achieving the regulatory traceability necessary for industrial deployment.

7. Conclusion

The integration of artificial intelligence (AI) with Raman spectroscopy represents a significant advancement in pharmaceutical analysis by combining molecular specificity with intelligent data processing. Throughout this review, it has been demonstrated that AI-assisted Raman spectroscopy enhances the accuracy and efficiency of pharmaceutical quality control, enabling rapid identification of active pharmaceutical ingredients, polymorphic forms, impurities, and counterfeit medicines. Machine learning algorithms, including Support Vector Machines, Random Forests, Artificial Neural Networks, and Convolutional Neural Networks, have shown excellent performance in interpreting complex Raman spectral data and supporting real-time decision-making.

The experimental findings further highlight the potential of AI-driven Raman analysis as a reliable Process Analytical Technology (PAT) tool for continuous pharmaceutical manufacturing. By enabling automated spectral interpretation and real-time process monitoring, this integrated approach can improve manufacturing efficiency, product consistency, and regulatory compliance while reducing analytical time and human intervention.Although challenges such as fluorescence interference, model interpretability, and the need for large, high-quality datasets remain, ongoing advances in explainable AI, portable Raman instrumentation, and intelligent manufacturing systems are expected to overcome these limitations. Overall, AI-assisted Raman spectroscopy provides a robust foundation for next-generation pharmaceutical manufacturing and is expected to play an increasingly important role in smart quality assurance, continuous process optimization, and the development of Industry 4.0 pharmaceutical production.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References

[1] He, Peiyu, Chunling Zhang, Yaosong Yang, Shuang Tang, Xixian Liu, Jin Yong, and Teng Peng. "Spectrum–effect relationships as an effective approach for quality control of natural products: a review." *Molecules* 28, no. pp. 1-26, 20 (2023): 7011. https://doi.org/10.3390/molecules28207011.

[2] Das, Ankita, and Parixit Prajapati. "Navigating pharmaceuticals: microfluidic devices in analytical and formulation sciences." Discover Chemistry 2, no. 1, pp. 1-23, (2025): 49. https://doi.org/10.1007/s44371-025-00133-y.

[3] Esmonde-White, Karen A., Maryann Cuellar, Carsten Uerpmann, Bruno Lenain, and Ian R. Lewis. "Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing." Analytical and bioanalytical chemistry 409, no. 3 (2017): pp. 637-649. https://doi.org/10.1007/s00216-016-9824-1.

[4] Ilchenko, Oleksii, Yurii Pilhun, Andrii Kutsyk, Denys Slobodianiuk, Yaman Goksel, Elodie Dumont, Lukas Vaut et al. "Optics miniaturization strategy for demanding Raman spectroscopy applications." Nature Communications 15, no. 1, pp. 1-14, (2024): 3049. https://doi.org/10.1038/s41467-024-47044-7.

[5] Elderderi, Suha, Franck Bonnier, Xavier Perse, Hugh J. Byrne, Florent Yvergnaux, Igor Chourpa, Abdalla A. Elbashir, and Emilie Munnier. "Label-Free Quantification of Nanoencapsulated Piperonyl Esters in Cosmetic Hydrogels Using Raman Spectroscopy." Pharmaceutics 15, no. 6, pp. 1-19, (2023): 1571. https://doi.org/10.3390/pharmaceutics15061571.

[6] Bufka, Jiří, Lenka Vaňková, Josef Sýkora, Věra Křížková, Jan Schwarz, and Petr Bouř. "Advancing Clinical Medicine with Raman Spectroscopy: Current Trends and Future Perspectives." Advanced Photonics Research 6, no. 12, pp. 1-24, (2025): 2500087. https://doi.org/10.1002/adpr.202500087.

[7] Shaikat, Mahdi Mubin, Venkata Gayatri Dhara, James K. Drennen, Guogang Dong, and Carl A. Anderson. "Enhancing analytical sensitivity in upstream bioprocess using time-gated Raman spectroscopy." Bioprocess and Biosystems Engineering 49, no. 2 (2026): pp. 385-401. https://doi.org/10.1007/s00449-025-03261-y.

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

[8] Liu, Yuan, Sitong Chen, Xiaomin Xiong, Zhenguo Wen, Long Zhao, Bo Xu, Qianjin Guo, Jianye Xia, and Jianfeng Pei. "Artificial intelligence guided Raman spectroscopy in biomedicine: Applications and prospects." Journal of Pharmaceutical Analysis, pp. 1-13, (2025): 101271. https://doi.org/10.1016/j.jpha.2025.101271.
[9] Lee, S., R. Park, and H. S. Jung. AI-enhanced surface-enhanced Raman scattering for accurate and sensitive biomedical sensing. Adv Intell Discov 2: e202500030, pp. 1-23, (2026). https://doi.org/10.1002/aidi.202500030.
[10] Coca-Lopez, Nicolas, Victor Alcolea-Rodriguez, Miguel A. Bañares, Sandor Brockhauser, Julien Gorenflot, Alex Henderson, Ron Hildebrandt et al. "Artificial intelligence-powered Raman spectroscopy through open science and FAIR principles." ACS nano 19, no. 44 (2025): 38189-38218. https://doi.org/10.1021/acsnano.5c09165.
[11] Chandra, A., Kumar, V., Garnaik, U. C., Dada, R., Qamar, I., Goel, V. K., & Agarwal, S. "Unveiling the molecular secrets: a comprehensive review of Raman spectroscopy in biological research." ACS omega, 9(51), (2024): 50049-50063. https://doi.org/10.1021/acsomega.4c00591.
[12] Li, Ying-Sing, and Jeffrey S. Church. "Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials." Journal of food and drug analysis 22, no. 1 (2014): 29-48. https://doi.org/10.1016/j.jfda.2014.01.003.
[13] Lin, Haonan, Hyeon Jeong Lee, Nathan Tague, Jean-Baptiste Lugagne, Cheng Zong, Fengyuan Deng, Jonghyeon Shin et al. "Microsecond fingerprint stimulated Raman spectroscopic imaging by ultrafast tuning and spatial-spectral learning." Nature communications 12, no. 1, pp. 1-12, (2021): 3052. https://doi.org/10.1038/s41467-021-23202-z.
[14] Fernández-Galiana, Álvaro, Olga Bibikova, Simon Vilms Pedersen, and Molly M. Stevens. "Fundamentals and applications of Raman-based techniques for the design and development of active biomedical materials." Advanced Materials 36, no. 43, pp 1-47, (2024): 2210807. https://doi.org/10.1002/adma.202210807.
[15] Lightner, Carin. "The Measurement of Raman Optical Activity." Ph.D. Thesis, ETH Zürich. (2022). https://doi.org/10.3929/ethz-b-000586562.
[16] Nicolau, Shelby T., Kenneth H. Jiang, and Adam J. Matzger. "Selecting a laser excitation source and sampling strategy for Raman spectroscopy." Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 343, pp. 50049-50063, (2025): 126588. https://doi.org/10.1016/j.saa.2025.126588.
[17] Kumar, Sumit, Zoltan Nagy, Gintaras V. Reklaitis, and Marcial Gonzalez. "Considerations in Raman Spectroscopy for Real-Time API Concentration Measurement at Tablet Press Feed Frame." Journal of Spectroscopy 2023, no. 1, pp. 1-16, (2023): 8631288. https://doi.org/10.1155/2023/8631288.
[18] Synytsya, Alla, Miluse Judexova, David Hoskovec, Michaela Miskovicova, and Lubos Petruzelka. "Raman spectroscopy at different excitation wavelengths (1064, 785 and 532 nm) as a tool for diagnosis of colon cancer." Journal of Raman Spectroscopy 45, no. 10 (2014): 903-911. https://doi.org/10.1002/jrs.4581.
[19] Liu, Yuan, Sitong Chen, Xiaomin Xiong, Zhenguo Wen, Long Zhao, Bo Xu, Qianjin Guo, Jianye Xia, and Jianfeng Pei. "Artificial intelligence guided Raman spectroscopy in biomedicine: Applications and prospects." Journal of Pharmaceutical Analysis, pp. 1-47, (2025): 101271. https://doi.org/10.1016/j.jpha.2025.101271.
[20] Jawaharsamuvel R., Sridevi A. R., Pallithara S. S. "A Comprehensive Review on Artificial Intelligence in Pharmaceutical Analytical Techniques: Advancements, Applications and Future Directions." International Journal of Pharmaceutical Sciences Review and Research, pp. 37-48 (2025): ISSN:0976 – 044X, 85(10). https://doi.org/10.47583/ijpsrr.2025.v85i10.006.
[21] Kalatzis, Dimitris, Alkmini Nega, and Yiannis Kiouvrekis. "Raman spectra classification of pharmaceutical compounds: a benchmark of machine learning models with SHAP-based explainability." Eng 6, no. 7, pp. 1-18, (2025): 145. https://doi.org/10.3390/eng6070145.
[22] Hardy, Mike, Pooja P. Kanade, Emma Buchan, Pola Goldberg Oppenheimer, Cillian McPolin, and Robert M. Bowman. "Miniaturisation of Raman spectroscopy systems: from benchtop to backpocket." Lab on a Chip 26, no. 7 (2026): 2112-2145. https://doi.org/10.1039/d5lc00836k.
[23] Mechram, Siti, Ieke Wulan Ayu, and Yulfita Farni. "Pretreatment method standard normal variate (SNV) and baseline shift correction (BSC) on the NIRS-based soil spectrum for rapid prediction of soil nitrogen content." In IOP Conference Series: Earth and Environmental Science, vol. 1290, no. 1, p. 012026. IOP Publishing, pp. 1-7, (2024). https://doi.org/10.1088/1755-1315/1290/1/012026.
[24] Bui, Dieu Tien, Paraskevas Tsangaratos, Viet-Tien Nguyen, Ngo Van Liem, and Phan Trong Trinh. "Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment." Catena 188, pp. 1-14, (2020): 104426. https://doi.org/10.1016/j.catena.2019.104426.
[25] Mozaffari, M. Hamed, and Li-Lin Tay. "A review of 1D convolutional neural networks toward unknown substance identification in portable Raman spectrometer." arXiv preprint arXiv:2006.10575, pp. 1-19, (2020).

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of **AJAPAS** and/or the editor(s). **AJAPAS** and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

 Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).