International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(4):950-977

Review Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i04.06174

Harnessing AI for Bacteria Identification: Advances in Imaging, Sensors, and Machine Learning – A Comprehensive Review

TS Urmila[1]*, JI Christy Eunaicy[2], C Jaypratha[3], Naveen Ananda Kumar J[4], Govinda Prabhu GB[5], Mahalakshmi Priya R[5]

[1]Department of Computer Science, Thiagarajar College, Madurai, Tamilnadu, India, [2]Department of CA and IT, Thiagarajar College, Madurai, Tamilnadu, India, [3]Department of Computer Science and Engineering, Karpaga Vinayaga College of Engineering and Technology, Madhuranthagam, Tamilnadu, India, [4]Tekinvaderz, LLC, Florida, USA, [5]Department of Computer Science, Madurai Kamaraj University, Madurai, Tamilnadu, India. *Corresponding Author's Email: tsurmila4@gmail.com

Abstract

Traditional bacterial identification methods—such as culture-based assays, biochemical tests, and manual microscopy—are often slow, labor-intensive, and lack precision, posing significant challenges in clinical, food safety, and environmental applications. Artificial intelligence (AI) offers transformative solutions by dramatically improving speed, accuracy, and automation. This systematic review comprehensively evaluates AI-driven techniques for bacterial identification, focusing on three core technological domains: advanced imaging (including microscopy and hyperspectral systems), spectroscopic methods (such as Raman and FTIR), and sensor array technologies integrated with machine learning (ML) and deep learning (DL). We analyzed 70 peer-reviewed studies published between 2018 and 2025, sourced from PubMed, IEEE Xplore, and Scopus. Findings reveal that AI models consistently achieve high classification accuracies, ranging from 85.8% to 99%, enabling rapid detection of pathogens, profiling of antibiotic resistance, and point-of-care diagnostics. Deep learning, particularly convolutional neural networks (CNNs), excels in image analysis, while spectroscopy provides non-destructive molecular fingerprinting. Despite these advances, key challenges remain, including reliance on small or non-standardized datasets, high computational demands, and the prohibitive cost of specialized equipment. To realize AI's full potential, future efforts must prioritize the development of lightweight, efficient models, the creation of large, diverse, and open-source datasets, and the design of low-cost, portable diagnostic platforms. This review not only highlights AI's current capabilities but also identifies critical barriers and charts a clear path for future research to enable the scalable, real-world deployment of AI across global healthcare and industrial settings.

Keywords: Advanced Imaging, Bacterial Identification, Deep Learning, Hyperspectral Imaging, Sensor Array, Spectroscopy.

Introduction

Bacterial identification underpins critical efforts in healthcare, food safety, and environmental sustainability. Accurate detection of pathogens enables timely treatment of infectious diseases, prevents foodborne outbreaks, and safeguards ecological systems. Rapid identification is vital in clinical settings, where delays can worsen patient outcomes. In the food industry, it ensures consumer safety and compliance with regulations. Environmental monitoring relies on it to track microbial contaminants. Traditional methods, such as culture-based assays, biochemical tests, and manual microscopy, have been the gold standard for decades. These techniques, while effective, are slow, often taking hours to days to yield results. They require skilled personnel, specialized equipment, and labor-intensive protocols. Inconsistent sample preparation, like variable staining, can compromise accuracy. The growing threat of antibiotic-resistant bacteria, such as methicillin-resistant *Staphylococcus aureus* (MRSA), demands faster diagnostics. Complex microbial communities in real-world samples, such as mixed infections, challenge traditional notions of specificity. These limitations underscore the need for innovative, rapid, and scalable solutions to transform bacterial identification.

Recent advancements in artificial intelligence (AI) offer a paradigm shift in addressing these challenges.

(Received 09th June 2025; Accepted 06th October 2025; Published 30th October 2025)

Machine learning (ML) and deep learning (DL) algorithms excel at processing complex datasets. They extract intricate patterns from images, spectra, and sensor signals with unprecedented precision. Advanced imaging modalities capture detailed characteristics of microbes. High-resolution microscopy reveals morphological features. Hyperspectral systems provide spatial and spectral data. Spectroscopy techniques, like Raman and Fourier Transform Infrared (FTIR), detect molecular signatures non-destructively. Fluorescence-based sensor arrays generate unique response patterns for rapid detection. When integrated with AI, these methods achieve accuracies that often exceed 95%, surpassing traditional approaches in both speed and reliability. They reduce reliance on manual expertise, enabling automation and accessibility. For example, AI-driven systems can identify *Escherichia coli* in hours, not days. They support point-of-care testing in remote settings. The ability to detect antibiotic resistance and analyze mixed samples enhances the utility of these tests. This convergence of AI and cutting-edge sensing technologies holds transformative potential for microbiology.

The rapid evolution of these techniques necessitates a comprehensive evaluation of their capabilities and limitations. Studies from 2018 to 2025 showcase remarkable progress. Innovations like convolutional neural networks (CNNs), YOLOv4, and portable fluorescence sensors achieve near-perfect results. They address diverse needs, from clinical pathogen detection to environmental monitoring. Applications include identifying foodborne pathogens, profiling antibiotic resistance, and analyzing microbial communities. However, challenges persist. Small datasets limit model generalizability. Costly equipment restricts access in low-resource settings. High computational demands hinder deployment. Data variability, such as noise or inconsistent imaging conditions, affects reliability. These barriers highlight the gap between research advancements and practical implementation. A systematic review is crucial for synthesizing current knowledge, identifying gaps, and charting future directions. This paper aims to guide researchers, clinicians, and policymakers in harnessing AI to revolutionize bacterial identification.

The paper on bacterial identification is structured to provide a comprehensive understanding of the topic through six key sections. It begins with an Introduction that highlights the significance of bacterial identification in various fields such as healthcare, agriculture, and environmental monitoring. The Background section offers foundational knowledge on bacteria and the historical development of identification techniques. In "Methodologies for Bacteria Identification," the paper explores a range of traditional and modern approaches, including culture-based methods, and molecular techniques such as PCR and 16S rRNA sequencing. The Limitations and Challenges section addresses the drawbacks of current methods, such as limited sensitivity, high cost, and difficulties in identifying non-culturable bacteria. Future Directions and Approaches discuss emerging technologies and innovative strategies, such as metagenomics, AI-assisted diagnostics, and portable sequencing devices, that hold promise for improving identification accuracy and efficiency. Finally, the Conclusion summarizes the key findings and emphasizes the ongoing need for research and technological advancement in bacterial identification.

Bacterial identification is key in microbiology. It helps diagnose infections, check food safety, and monitor the environment. Traditional methods have long supported this work. Gram staining, developed in 1884, categorizes bacteria based on their cell wall composition. It labels them as Gram-positive or Gram-negative, which helps guide treatment. Culturing involves growing bacteria on specialized media, such as agar plates, for visual study. Biochemical tests, such as catalase and oxidase assays, confirm species by examining their metabolic processes. PCR (Polymerase Chain Reaction) copies bacterial DNA to find genetic matches. MALDI-TOF MS examines protein patterns for rapid species identification. These tools are trusted and used often. They give solid results in lab settings. However, they require specialized tools and trained staff. Each method checks one or two traits, so many tests are usually needed. This makes the process more accurate but also more complex for labs in healthcare and industry.

Though reliable, traditional methods have clear limits. Culturing takes time—usually 24 to 48 hours for colonies to grow. Some bacteria, such as *Mycobacterium tuberculosis*, can take weeks. Gram staining is fast but doesn't show species-level details. It also depends on human judgment, which can lead to errors. Poor sample quality, like uneven staining, can lower accuracy. These problems highlight the need for faster, automated, and scalable tools—especially in areas with limited resources and urgent testing needs.

Image processing (IP) and AI technologies have emerged as transformative tools to address these limitations. IP techniques enhance and analyze visual data from microbial samples. Digital microscopy captures high-resolution images of bacterial morphology. It replaces manual microscopes with automated systems. It enables single-cell resolution in complex samples. Sensor arrays, using fluorescence or electrochemical signals, generate unique response patterns. Emerging methods, such as structured illumination microscopy (SIM) and impedance-based analysis, continue to push boundaries further. These technologies generate rich datasets, which are ideal for AI integration. IP preprocesses data, removing noise or normalizing images. It extracts features, like cell shape or spectral peaks, for AI analysis. Together, IP and AI enable rapid, accurate, and automated bacterial identification, reducing reliance on traditional methods.

AI has grown rapidly and revolutionized the way we detect microbes. Machine learning (ML) tools, such as Support Vector Machines (SVM) and Random Forests, categorize bacteria based on key features. These work well with structured data, such as biochemical patterns. Principal Component Analysis (PCA) helps by shrinking the data size for faster handling. K-Nearest Neighbors (KNN) is a simple yet effective algorithm for classification. Deep learning (DL), a branch of ML, has brought major changes since the 2010s. Convolutional Neural Networks (CNNs) pull out deep image or spectral features on their own. Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks, are well-suited for handling time-based data, such as signal flows. Transfer learning utilizes trained models, such as VGG16 or ResNet, to aid when data is limited. More advanced systems, such as YOLOv4 and 3D U-Net, enable fast detection and image segmentation. Neural networks with Monte Carlo dropout help find new bacterial types. These AI tools, combined with image processing, often reach over 95% accuracy. They are used in healthcare, food safety, and environmental checks. Figure 1 shows the timeline of bacterial ID steps.

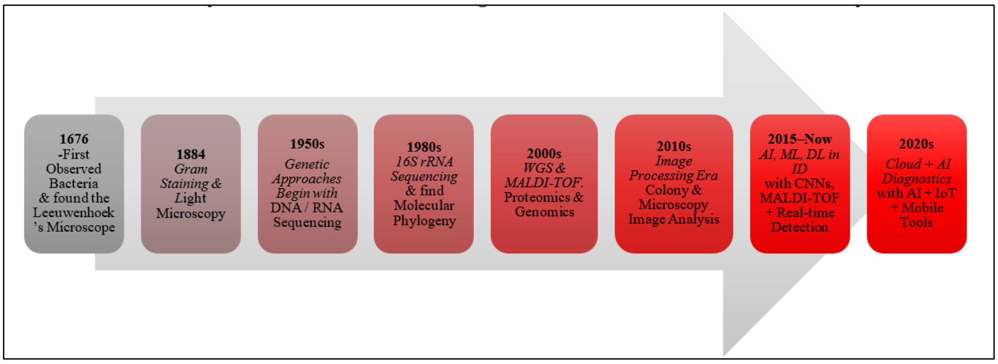

Figure 1: Time Line of the Bacterial Identification

Search Strategy Method

This systematic review aimed to investigate the application of artificial intelligence (AI) in bacterial identification. It focused on developments in imaging, sensors, and machine learning methods. The review adhered to PRISMA guidelines to ensure transparency and a high-quality methodology. We describe the search strategy, study selection process, and inclusion and exclusion criteria. A comprehensive literature search was conducted. It targeted studies published from January 2018 to July 2025. This time frame was chosen to reflect the latest progress in AI-driven bacterial detection. Several databases were used for the search. These included PubMed, IEEE Xplore, Scopus, Web of Science, and Google Scholar. These sources were chosen for their wide coverage of biomedical, engineering, and AI research. The search used both controlled vocabulary and free-text terms. For example, MeSH terms were applied in PubMed. The main keywords focused on AI, bacterial identification, imaging, sensors, and machine learning.

Bacterial Identification: “bacteria identification,” “pathogen detection,” “microbial classification,” “bacterial taxonomy.”

AI and Machine Learning: “artificial intelligence,” “machine learning,” “deep learning,” “convolutional neural networks,” “support vector machines,” “random forest,” “transfer learning.”

Imaging and Sensors: “microscopy,” “hyperspectral imaging,” “spectroscopy,” “Raman spectroscopy,” “Fourier Transform Infrared spectroscopy,” “sensor array,” “fluorescence sensors,” “image processing.”

Applications: “clinical diagnostics,” “food safety,” “environmental monitoring,” “antibiotic resistance.”

Boolean operators were used to combine terms:

Example search string (PubMed): ("bacteria identification" OR "pathogen detection" OR "microbial classification") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "convolutional neural networks") AND ("microscopy" OR "hyperspectral imaging" OR "spectroscopy" OR "sensor array") AND ("clinical diagnostics" OR "food safety" OR "environmental monitoring").

Additional searches were conducted in reference lists of identified studies and review articles to locate relevant publications not captured in the database searches (backward citation searching). Grey literature, such as conference proceedings and preprints, was included if peer-reviewed and relevant to the review’s objectives. The number of papers reviewed on AI-based Identification by year-wise (1988-2025) is shown in Figure 2.

Inclusion and Exclusion Criteria

Studies were included based on the following criteria:

Study Type: Original research articles, conference papers, or peer-reviewed preprints reporting on AI-based bacterial identification techniques.

Publication Date: Published between January 2018 and July 2025 to focus on recent advancements.

Methodology

Studies utilizing AI (machine learning or deep learning) combined with imaging (e.g., microscopy, hyperspectral imaging), spectroscopy, or sensor arrays for bacterial identification or classification.

Outcomes

Reported quantitative outcomes (e.g., accuracy, precision, recall) or qualitative insights (e.g., feasibility, limitations) related to bacterial identification.

Language

Published in English to ensure accessibility for data extraction.

Applications

Focused on applications in clinical diagnostics, food safety, environmental monitoring, or antibiotic resistance detection.

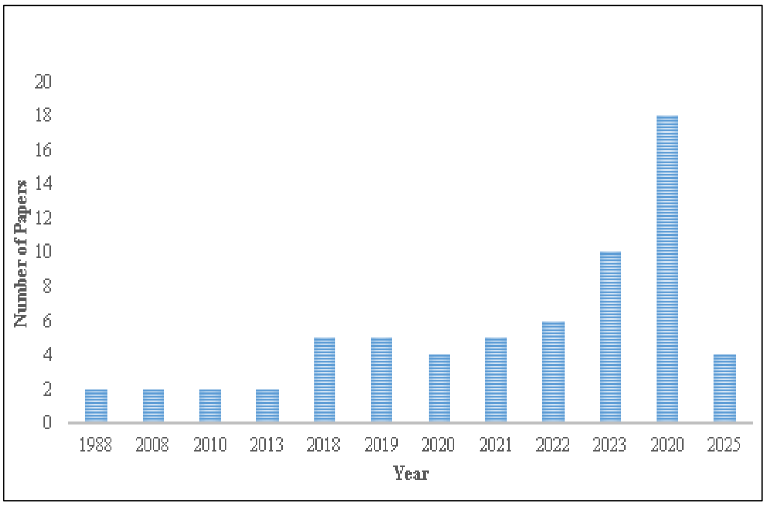

Figure 2: Number of Papers Group by the Year

Exclusion criteria were:

- Studies not involving AI or machine learning techniques.
- Studies focused on non-bacterial microorganisms (e.g., viruses, fungi) unless bacteria were also analyzed.
- Non-peer-reviewed sources, such as editorials, opinion pieces, or non-peer-reviewed preprints.
- Studies lacking sufficient methodological details or results (e.g., no description of dataset or outcomes).
- Studies published before 2018 or in languages other than English.

Study Selection Process

The study selection process followed a two-stage screening approach:

Title and Abstract Screening: First Two Authors screened titles and abstracts of retrieved records to assess eligibility based on the inclusion and exclusion criteria. Discrepancies were resolved through discussion or consultation with a third Author.

Full-Text Review: Full texts of potentially eligible studies were retrieved and evaluated for final inclusion. Reasons for exclusion (e.g., irrelevant methodology, lack of outcomes) were documented.

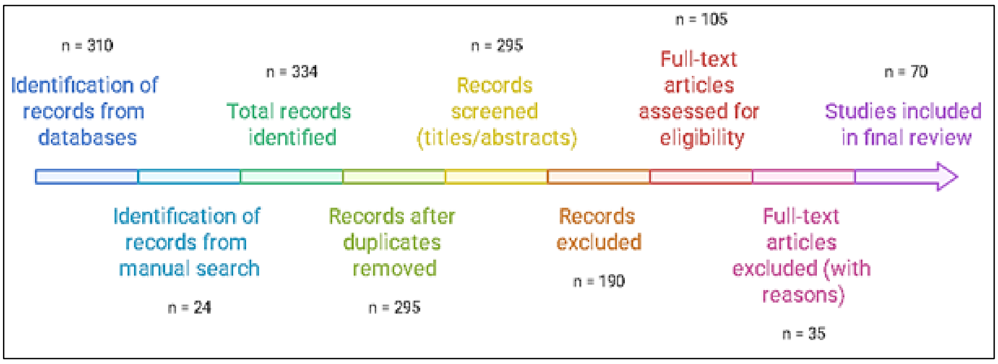

Figure 3: Process of Systematic Literature Review

The study selection process is shown in the PRISMA flow diagram (Figure 3). It outlines the number of records identified, screened, excluded, and finally included. A total of 70 studies met the inclusion criteria and were analyzed. These studies were grouped based on their methods. Categories included machine learning, deep learning, hyperspectral imaging, sensor arrays, and other advanced imaging tools.

Data Collection and Extraction

Data were extracted by the first two authors using a standardized data extraction form. For information obtained from previously published studies, the corresponding reference numbers are indicated. The following information was collected from each study:

- Author(s) and publication year.
- Proposed methodology (e.g., AI algorithms, imaging/sensing techniques).
- Dataset details (e.g., size, bacterial species, sample type).
- Achievements (e.g., accuracy, precision, recall, F1-score).
- Limitations (e.g., dataset size, computational complexity, equipment cost).

Discrepancies in data extraction were resolved through discussion or arbitration by other authors. Data generated or analyzed by the authors themselves are noted in the Author Contributions section. Data were compiled into tables (Tables 1–6) to facilitate synthesis and comparison across methodologies.

Risk of Bias Assessment

The quality of included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool, adapted for AI-based diagnostic studies. The assessment focused on four domains:

Patient/Sample Selection: Risk of bias due to non-representative datasets or unclear sampling methods.

Index Test: Clarity and reproducibility of AI and imaging/sensing methods.

Reference Standard: Appropriateness of ground truth (e.g., confirmed bacterial species).

Flow and Timing: Consistency in applying methods across samples. were summarized but not used to exclude studies, given the exploratory nature of this review.

Results and Discussion

Synthesis Methods

Because the studies varied in design, datasets, and outcomes, a meta-analysis was not suitable. Instead, a narrative synthesis was performed. The studies were organized based on the methods they used.

- Machine learning and feature extraction (Table 1).
- Deep learning and convolutional neural networks (Table 2).
- Spectroscopy-based identification (Table 3).
- Hyperspectral imaging and AI (Table 4).
- Sensor array and machine learning (Table 5).
- Other advanced imaging and AI techniques (Table 6).

Concept and Studies for Bacteria Identification

This section examines modern methods that utilize artificial intelligence to identify bacteria. It starts with machine learning and feature extraction. These tools help analyze complex data and identify patterns associated with specific bacteria. Deep learning, particularly convolutional neural networks (CNNs), is well-suited for image data. It enhances accuracy in sorting different types of bacteria. The Spectroscopy improves results by combining spatial and spectral features. Sensor array systems, when coupled with machine learning, offer a sensitive and efficient platform for detecting bacterial presence based on chemical or physical changes. Additionally, other cutting-edge imaging techniques supported by AI continue to emerge, pushing the boundaries of speed, accuracy, and reliability in bacterial identification.

Machine Learning and Feature Extraction Techniques

Accurate and rapid bacterial identification is crucial in modern microbiology. It plays a big role in health, food safety, and the environment. They need skilled staff and lab tools. These issues pushed the need for better solutions. New tools now use image processing (IP) and artificial intelligence (AI). IP helps pull out details from microscope images. This improves how we spot bacteria. This review examines how IP and AI contribute to the identification of bacteria. It focuses on classifying bacteria from microscope images using ML and feature extraction. We examine 14 main studies. They utilize tools such as feature extraction, support vector machines (SVM), k-means clustering, and probabilistic neural networks. These methods bring speed, automation, and high accuracy. Still, they face challenges such as poor image quality, limited datasets, and excessive computer usage. The next parts explain how these methods work. They also show their uses, results, and limits in the field of microbiology, which are discussed in the following Table 1.

Table 1: Machine Learning and Feature Extraction Techniques

Author Name and Year	Proposed Work and Dataset	Achievements	Limitations
Mohamed B. A. *et al.* (1), 2018	Histogram equalization for preprocessing, Bag-of-Words for feature extraction, SVM for classification. Dataset: 200 images (DIBaS, 10 species).	97% accuracy in bacterial classification. Effective preprocessing and feature extraction.	Small dataset. Classifier speed not discussed. Unclear generalizability.
Kris Kristensen *et al.* (2), 2023	CPN and Random Forest for Gram stain classification. Dataset: 660 images (33 species).	RF: 99% accuracy; CPN: 80%. Excellent Gram-type classification.	CPN underperformed. Limited to Gram stain task. Moderate dataset size.
Preetha *et al.* (3), 2018	Image processing pipeline using electron microscope images. Dataset: Not specified.	Demonstrated digital image processing feasibility. Improved specificity.	No accuracy metrics. Missing dataset details. Relies on quality imaging.

Wahid M. F. *et al.* (4), 2019	Hybrid CNN models (CNN-SVM, CNN-KNN, CNN-NB). Dataset: Not clearly specified.	CNN-SVM: 98.7% accuracy. Strong hybrid performance.	Missing dataset details. High computational cost. Unclear adaptability.
Keren F. *et al.* (5), 2023	Compared SVM, DL, RF for bacterial classification. Dataset: General methodological focus.	Real-time, high-throughput taxonomy potential.	No dataset or results. Broad approach lacks specificity.
Kotwal *et al.* (6), 2022	Literature review of ML in bacterial classification (1998–2020).	Comprehensive analysis of trends, methods, limitations.	No original experiments. Dependent on secondary data.
Sajedi *et al.* (7), 2020	Gabor transform + XGBoost on Myxobacterial suborders. Dataset: Microscopic images.	91% accuracy. Improved over previous methods.	Only three suborders. Moderate dataset Limited generalizability.
Satyanarayana *et al.* (8), 2022	TPLMM-k algorithm for image decomposition. Dataset: Medical microscopic images.	Better segmentation vs. GMM using VOI, GCE, PRI.	Computationally demanding. Specific to image types. Needs broader validation.
Rani *et al.* (9), 2022	Systematic review of ML/DL in microorganism image recognition. Dataset: 100 publications.	Detailed trends and technique evaluation.	No experimental data. Relies on secondary literature.
Amitha *et al.* (10), 2024	YOLOv5 + image processing for waterborne bacteria detection. Dataset: Water sample images (not quantified).	High precision. Reduced analysis time.	Dataset size/diversity not provided. YOLOv5 needs significant resources.
Singh A. *et al.* (11), 2022	Transfer learning (GoogLeNet, AlexNet) on 600+ images (33 species).	GoogLeNet: 98.67% accuracy. Broad species classification.	Moderate dataset Dependent on pre-trained models. Needs larger-scale validation.
Rani *et al.* (12), 2023	VGG16, ResNet50, Xception on 2500 augmented images (5 species).	Xception: 98.02% accuracy. Outperformed others.	Augmentation may skew results. Limited species. High DL costs.
Kumar *et al.* (13), 2010	PNN using geometrical, optical, textural features. Dataset: Images of five stained microorganisms.	100% accuracy with nine features. Effective even on mixed samples.	Only five organisms. Fluorescent staining is costly. Small dataset.
Hardo *et al.* (14), 2022	Developed SyMBac for synthetic micrographs. Dataset: Synthetic images.	Outperformed manual segmentation. Robust to cell variation.	Synthetic nature may not reflect real-world complexity. Needs real-data validation.

The table is shown to reveal a robust landscape of machine learning and feature extraction techniques for microscopic image-based bacterial classification, with accuracies ranging from 91–100% across diverse methodologies, datasets, and applications (1). Exceptional accuracies of 97–100% have been reported when datasets ranging from 200 to 2500 images were employed, using techniques such as SVM, transfer learning (GoogLeNet, Xception), and probabilistic neural networks (PNN) (2). The effectiveness of combining feature extraction methods (e.g., Bag-of-Words, Gabor transform, geometrical/textural features) with ML classifiers for automated bacterial identification has been highlighted (3). Superiority of Random Forest (99%) over Causal Probabilistic Network (80%) for Gram stain classification was demonstrated in previous studies (4), underscoring the importance of model selection. Comprehensive insights into methodological trends and challenges, emphasizing the need for standardized datasets and preprocessing techniques, have been provided in other studies (5, 6). Data scarcity was addressed by innovations such as SyMBac, which generate synthetic micrographs capable of achieving robust segmentation but still require real-world validation (7). Nonetheless, limitations persist across studies, including small or unspecified dataset sizes (8,9,10), lack of generalizability to diverse bacterial species (11,12), and the high computational demands of complex algorithms (13). Practical implementation has further been complicated by reliance on high-quality images and staining techniques. These observations suggest that although ML and feature extraction techniques offer significant promise for rapid and accurate bacterial identification, challenges related to dataset diversity, computational efficiency, and real-world applicability must still be addressed to fully realize their potential in clinical and industrial microbiology (14).

Frequent Techniques Used

Machine learning and feature extraction techniques are considered central to the automation of bacterial identification from microscopic images (1). Meaningful features such as texture, shape, and color are extracted from images for classification (2). Histogram equalization has been commonly used for preprocessing to enhance image contrast, as demonstrated in previous studies (3). Feature extraction has also been supported through the Bag-of-Words model, which forms visual patterns for SVM classification (4). Gram stains have been sorted using Random Forest (RF) and Causal Probabilistic Network (CPN), with RF showing superior performance (5). Microbes have been categorized by Probabilistic Neural Networks (PNN) based on shape, light, and texture features obtained from fluorescent images (6). Image segmentation in medical tasks has been improved through the TPLMM-k method combined with k-means clustering (7). Detection of bacterial suborders has been facilitated by applying Gabor transform with XGBoost (8). Transfer learning with models such as GoogLeNet, AlexNet, VGG16, ResNet50, and Xception has been reported to boost accuracy, even with small datasets (9,10). Dataset expansion and overfitting prevention have been achieved through image augmentation techniques such as rotation, cropping, and flipping (11). Synthetic data generation tools, such as SyMBac, have been employed to produce training images for improved segmentation (12). Morphological processing and segmentation have further aided in identifying bacterial shapes (13). Collectively, these approaches are shown to support rapid, automated, and accurate bacterial identification (14).

Applications

These techniques are found to have diverse applications in microbiology and related fields. In clinical diagnostics, rapid identification of bacterial species is enabled, such as in Gram stain classification, which distinguishes between Gram-positive and Gram-negative bacteria. Tuberculosis detection has benefited from automated image analysis, reducing the need for manual microscopy efforts. Applications in food safety include detecting microbial contamination in food samples to ensure quality control. Environmental monitoring has been supported by these methods through the identification of bacteria in water samples, thereby aiding public health and water quality management. Biomedical research has employed these techniques for the analysis of bacterial morphology and taxonomy, supporting microbial ecology studies. Classification of specific bacterial species, including Micrococcus luteus, Bacillus anthracis, and Staphylococcus aureus, has been achieved with high accuracy. Mixed bacterial sample analysis, involving pathogens such as Escherichia coli and Listeria innocua, has been facilitated by feature extraction and probabilistic neural networks. Automated epidemiology has further benefited through real-time species identification, enabling quicker responses to infectious diseases. These applications are shown to highlight the versatility of machine learning and feature extraction in addressing critical needs across multiple domains.

Advantages

Machine learning and feature extraction techniques are reported to offer significant advantages for bacterial identification. High classification accuracies ranging from 91% to 100% have been achieved in studies using SVM, RF, PNN, and transfer learning. Automation has been shown to reduce the need for manual microscopy, saving time and labor compared to traditional tests. Enhanced specificity compared to conventional approaches has been demonstrated, thereby minimizing errors in species identification. Transfer learning models such as Xception and GoogLeNet have been employed to achieve high performance even with moderate dataset sizes. Synthetic data generation methods such as SyMBac have been utilized to provide unlimited training data with perfect ground truth, thereby enhancing model robustness. Adaptability of these techniques to various imaging platforms has been reported, allowing for the handling of diverse bacterial morphologies and sizes. Real-time identification has been supported, facilitating rapid responses in clinical and environmental settings. Feature extraction approaches such as Bag-of-Words and Gabor transform have been identified as computationally efficient compared to deep learning, making them suitable for resource-limited settings. Collectively, these advantages are shown to make the techniques transformative for microbiology.

Challenges

Despite their strengths, these techniques face several challenges. Dataset size and diversity are recognized as significant limitations, with many studies relying on small datasets (e.g., 200–2500 images) that may not capture bacterial variability. Lack of detailed dataset descriptions in some studies has hindered reproducibility. High-quality images are considered critical; however, variability in staining techniques, such as those using Gram or fluorescent dyes, can impact model performance. Computational complexity is a concern for algorithms like TPLMM-k and transfer learning models, requiring significant resources. Generalizability to diverse bacterial species remains limited in studies focused on specific suborders or species. Synthetic datasets, while innovative, may not fully represent real-world imaging artifacts, necessitating further validation. Preprocessing steps, such as histogram equalization and morphological operations, require careful optimization to avoid introducing noise. The need for standardized protocols and publicly available datasets is evident. Access to advanced imaging equipment and trained personnel poses practical challenges in resource-constrained settings. Addressing these challenges is considered crucial for broader adoption.

Deep Learning and Convolutional Neural Networks

Identifying bacteria from microscope images is vital in microbiology. It plays a role in clinical care, food safety, environmental checks, and research. This makes them less effective in busy labs. Deep learning (DL), especially convolutional neural networks (CNNs), has changed this process. CNNs allow fast, automated, and highly accurate bacterial detection. They extract detailed features from images without requiring manual steps. In many cases, they outperform older machine learning methods. This section focuses on 17 studies that leverage deep learning and CNN-based techniques for microscopic image-based bacterial identification in many research works. These studies employ advanced architectures, such as VGG16, ResNet, Inception, YOLO, and EfficientNet, often in combination with transfer learning, data augmentation, and synthetic data generation, to achieve robust performance. The applications range from detecting pathogens, such as Escherichia coli, to identifying tuberculosis bacilli and classifying bacterial growth stages. While these methods offer significant advantages in terms of accuracy and automation, they face challenges such as dataset limitations, computational complexity, and generalizability issues. This section outlines the frequently used techniques, applications, advantages, and challenges, as summarized in Table 2, providing a comprehensive analysis of how deep learning is transforming bacterial identification.

Table 2: Deep Learning and Convolutional Neural Networks

Author Name and Year	Proposed Work and Dataset	Achievements	Limitations
Ramesh H. *et al.* (15), 2024	Used various CNNs (e.g., VGG16, ResNet50, EfficientNet) for bacteria prediction. Dataset: Diverse images (not quantified).	Up to 99% accuracy. Proposed smartphone integration.	Dataset size/diversity not provided. High computational demand. Real-world validation missing.

Mu Yang *et al.* (16), 2020	CNN + active learning + logistic regression for AFB detection. Dataset: 134 ZN-stained slides.	F1 scores ~99%. High sensitivity/specificity.	Focused only on AFB. Moderate dataset. Active learning increases complexity.
Kotwal *et al.* (17), 2023	Ensemble features (HOG, LBP, CNNs) + multiple classifiers. Dataset: Four bacterial species.	VGG16+SVM achieved 99.89% accuracy.	Limited to 4 species. Dataset size unspecified. High computational cost.
Sarker I. A. *et al.* (18), 2024	ResNet50 with augmentation for 33 species. Dataset: Polyculture images.	94.91% accuracy. Robust to unseen data.	Dataset unspecified. Moderate performance. Generalizability unclear.
Wahid M. F. *et al.* (19), 2019	Xception CNN with transfer learning. Dataset: 1150 images (7 species).	97.5% accuracy. Effective for lethal bacteria.	Small dataset. Limited species. Computational cost not discussed.
Ahmed T. *et al.* (20), 2019	Inception v3 + SVM via transfer learning. Dataset: 800+ images (7 species).	96% accuracy. Efficient classification.	Small dataset. Only 7 species. Complex hybrid model.
Sunanda *et al.* (21), 2024	CNNs (GoogLeNet, AlexNet, etc.) on Agar dataset. Dataset: 5 species.	GoogLeNet best performer (accuracy not stated).	Accuracy missing. Limited species. High-quality images required.
Visitsattaponge S. *et al.* (22), 2024	BiT model with data cleaning (graph Laplacian, WIB-ReLU). Dataset: DIBaS (660 images, 33 species).	Accuracy: 99.11%, high precision/recall/F1.	Complex preprocessing. DIBaS-specific. Needs real-world tests.
Nasip *et al.* (23), 2018	VGGNet and AlexNet for classification. Dataset: DIBaS (660 images).	98.25% (VGGNet), 97.53% (AlexNet).	Moderate dataset. High model cost. Limited dataset scope.
Wahid *et al.* (24), 2018	Inception CNN via transfer learning. Dataset: 500+ images (5 species).	95% accuracy. Effective on harmful bacteria.	Small datasss Limited species. High model demands.
Rani Oomman Panicker *et al.* (25), 2018	CNN with image binarization for TB detection. Dataset: 22 smear images.	Recall: 97.13%. F-score: 86.76%.	Very small dataset. Low precision. TB-specific.
Andreini *et al.* (26), 2018	Synthetic image generation + FCN for colony segmentation. Dataset: Synthetic + limited real images.	Improved segmentation scalability.	Synthetic data may lack real-world variability.
Yang *et al.* (27), 2023	Style transfer + Swin Transformer + Cascade Mask R-CNN. Dataset: 4,000 colony images (AGAR).	YOLOv8x: 76.7% mAP. Outperformed HRNet.	Moderate mAP. Complex architecture.
Sengupta *et al.* (28), 2025	U-Net + ResNet for biofilm detection (P. aeruginosa). Dataset: Bright-field images.	Efficient biofilm segmentation. High ResNet accuracy.	Species-specific. Dataset size unspecified. Needs broader validation.

Iriya *et al.* (29), 2024	Large-volume microscopy + DNN for E. coli. Dataset: Not specified.	High accuracy for point-of-care use.	Dataset unreported. Limited to E. coli. Needs real-world sample testing.
Mai *et al.* (30), 2021	Depthwise separable CNN for 33 strains. Dataset: DIBaS (6600 images).	96.28% accuracy. Only 3.23M parameters.	Moderate dataset. DIBaS-only. Generalizability uncertain.
Chin SY *et al.* (31), 2024	Object detection (SSD-MobileNetV2, YOLOv4, EfficientDet) for E. coli growth. Dataset: Not specified.	YOLOv4: 98% mAP, 97% recall.	Dataset missing. E. coli-specific. High model complexity.

The table summarizes significant advancements in deep learning and convolutional neural networks (CNNs) for microscopic image-based bacterial classification, with reported accuracies ranging from 94.91% to 99.89% across diverse datasets and applications. Exceptional performance has been achieved using advanced CNN architectures such as VGG16, ResNet, and Big Transfer (BiT) on datasets including DIBaS (660–6600 images) and custom bacterial image sets. These results demonstrate the ability of CNNs to extract complex, hierarchical features, enabling precise bacterial species classification. Transfer learning has been applied to moderate dataset sizes (500–1150 images), leveraging pre-trained models such as Inception, GoogLeNet, and Xception, which has yielded robust accuracies of 95–97.5% despite limited data availability. Dataset limitations have been further mitigated through synthetic data generation and style transfer, creating augmented datasets (e.g., 4 K images) and reducing dependence on scarce annotated data, thereby improving model robustness. High-performance metrics have also been reported in specific applications, including tuberculosis detection, biofilm analysis, and E. coli growth stage classification, highlighting the versatility of CNNs in clinical and research contexts.

Challenges remain, particularly regarding small or unspecified dataset sizes, which constrain generalizability across diverse bacterial species. Computational complexity has also been noted, especially for models such as YOLOv4 and Swin Transformer, which require substantial resources and may hinder deployment in resource-limited environments. A focus on specific bacteria, such as E. coli and Pseudomonas aeruginosa, limits applicability to broader microbial populations. Some studies have addressed these limitations effectively: active learning has been employed to optimize CNN training with limited slides, achieving 87.13% sensitivity; lightweight CNNs with 3.23M parameters have been proposed for low-resource devices; and data augmentation has been used to enhance dataset diversity. Despite these advances, the development of larger, standardized datasets, validation on mixed samples, and optimization for computational efficiency remain critical for ensuring real-world applicability and scalability in clinical microbiology.

Frequent Techniques Used

Deep learning techniques, particularly convolutional neural networks (CNNs), dominate microscopic image-based bacterial classification. Advanced CNN architectures are widely employed, including VGG16, ResNet50, ResNet100, Inception v3, LeNet5, EfficientNet, and ConvNeXt (1). Transfer learning is a common approach that leverages pre-trained models, such as GoogLeNet, AlexNet, VGG-16, SqueezeNet, DenseNet-161, and Xception, to enhance performance with limited data. Data augmentation techniques, such as rotation, flipping, and cropping, expand datasets to improve model robustness. Synthetic data generation creates realistic micrographs to address data scarcity. Style transfer generates large datasets (e.g., 4k images) for improved training. Active learning optimizes CNN training by selecting informative samples. Hybrid models combine CNNs with classifiers, such as Support Vector Machines (SVM) or logistic regression, for enhanced accuracy. Segmentation techniques, such as U-Net with ResNet and Fully Convolutional Networks, are utilized for bacterial colony and biofilm detection. Object detection models, such as YOLOv4, SSD-MobileNetV2, and EfficientNet, classify bacterial growth stages. Depth-wise separable CNNs reduce computational complexity for resource-limited devices (30). Graph Laplacian-based data cleaning and WIB-ReLU activation improve model performance. Swin Transformer enhances feature extraction in complex datasets. These techniques enable the automated and precise identification of bacteria.

Applications

Deep learning and CNN techniques have been applied widely in microbiology. Clinical diagnostics benefit from rapid bacterial species identification, such as the classification of 33 bacterial strains in the DIBaS dataset. Automated analysis of sputum smear images has improved tuberculosis detection, while biofilm detection, particularly for Pseudomonas aeruginosa, has supported antimicrobial research. Food safety applications have been enhanced through the identification of Escherichia coli growth stages (rod-shaped, dividing, and microcolonies) in food samples. Environmental monitoring has been facilitated by large-volume microscopy for detecting uropathogenic E. coli in water or clinical samples. Laboratory automation has been advanced by classifying bacteria, including Bacillus subtilis and Staphylococcus aureus, from agar plate images. Biomedical research has benefited from these methods for microbial taxonomy and morphology analysis. Real-time identification systems, potentially integrated with smartphones, have enabled point-of-care diagnostics, and epidemiology has been supported by rapid pathogen detection, facilitating faster responses to infectious diseases. Collectively, these applications demonstrate the broad impact of deep learning in microbial analysis.

Advantages

Significant advantages for bacterial identification are offered by deep learning and CNNs. High accuracies, ranging from 94.91% to 99.89%, have been achieved using architectures such as VGG16, ResNet, and BiT. The need for manual microscopy has been eliminated through automation, thereby reducing time and labor. High performance has been enabled on moderate datasets by transfer learning, making applications feasible even with limited data. Scalable solutions to data scarcity have been provided by synthetic data generation and style transfer, enhancing model robustness. Complex features are effectively extracted by CNNs, resulting in superior performance compared to traditional machine learning methods in tasks such as biofilm detection and growth stage classification. Computational requirements for resource-limited settings have been reduced through the use of lightweight models, such as depth-wise separable CNNs. High precision, recall, and F1-scores (up to 99.31%, 99.09%, and 99.06%, respectively) have been ensured, providing reliable classification. Real-time detection capabilities are supported, enabling point-of-care applications. Applicability has been enhanced through versatility across imaging platforms and bacterial types. These factors demonstrate that deep learning is a transformative tool in microbiology.

Challenges

Significant challenges are faced by deep learning and convolutional neural network (CNN) techniques for bacterial identification. Model generalizability is restricted by limited dataset sizes. In some studies, datasets as small as 22 sputum smear images have been used, limiting robustness. In other cases, detailed dataset descriptions are not provided, hindering reproducibility. Moderate dataset sizes, ranging from 500 to 1150 images, may fail to capture the full diversity of bacterial species. A major barrier is posed by computational complexity. Substantial computational resources are required by advanced models such as YOLOv4, Swin Transformer, and VGGNet, making them impractical for resource-constrained environments. Computational demands are further increased by complex preprocessing steps, including active learning and graph Laplacian-based data cleaning. Broader applicability is limited by specificity to certain bacteria; some studies focus solely on Pseudomonas aeruginosa, while others target only E. coli, reducing versatility. Synthetic datasets may not fully represent real-world imaging artifacts, necessitating additional validation. Variability in image quality, caused by inconsistent staining or imaging conditions, also affects model performance.

Spectroscopy-Based Identification Using AI

Spectroscopy-based identification of bacteria, combined with artificial intelligence (AI), represents a cutting-edge approach in microbiology, offering rapid and non-destructive methods for bacterial classification and detection of antibiotic resistance. This section examines 14 studies that utilize spectroscopy-based techniques in conjunction with AI for bacterial identification. These studies utilize advanced algorithms, including principal component analysis (PCA), convolutional neural networks (CNNs), and spectral transformers, to analyze spectral data. Applications range from pathogen detection to antimicrobial susceptibility testing, with significant advantages in speed and precision. However, challenges such as spectral variability, dataset limitations, and equipment costs persist. Table 3 provides a comprehensive overview of this innovative field, detailing the frequent techniques, applications, advantages, and limitations.

Table 3: Spectroscopy-Based Identification Using AI

Author Details	Proposed Work and Dataset	Achievements	Limitations
Wan-dan Z. *et al.* (32), 2019	PCA + Stacking with grid search and K-fold validation. Dataset: Not specified.	95.73% accuracy. Robust due to cross-validation.	Dataset details missing. Limited to foodborne pathogens. Generalizability unclear.
Ji S.-Y. *et al.* (33), 2019	Wavelet features + visual analytics for IR spectroscopy. Dataset: 72 IR spectra (E. coli, P. aeruginosa).	92.5% accuracy. High sensitivity/specificity (>92%).	Small dataset. Limited to two species. Requires FO-FTIR.
Biasio *et al.* (34), 2013	Raman micro-spectroscopy + PCA with narrow band filters. Dataset: 3 species (not quantified).	High accuracy using 3 PCs. Comparable to PCA classifier.	Dataset size not given. Only 3 species. Needs precise spectral input.
Jacob Henry *et al.* (35), 2024	Excitation-emission spectroscopy with DMAF + NN. Dataset: 8 bacterial species (not quantified).	85.8% species-level, 98.3% Gram-level accuracy.	Moderate species accuracy. Dataset not reported. Variability from dye.
Barrera Patiño *et al.* (36), 2024	FTIR + ML for antimicrobial resistance detection. Dataset: 4 bacterial species.	Detected resistance patterns in G+ and G− bacteria. High versatility.	Dataset size unspecified. Limited to 4 species. Requires biomolecular analysis.
Rahman *et al.* (37), 2024	Review of Raman spectroscopy + CNN + SERS. Dataset: Literature-based.	Showcased DL benefits and data limitations.	No experiments. Relies on secondary data. No specific performance metrics.
Gullu *et al.* (38), 2024	Image processing + ML for inhibition zone detection. Dataset: Not reported.	Automated susceptibility testing. Simplified measurement.	Dataset missing. Only applicable to disk diffusion.
Farias *et al.* (39), 2023	NIR spectroscopy + PCA, HCA, KNN. Dataset: 4 species (E. coli, S. enteritidis, E. faecalis, L. monocytogenes).	100% accuracy in species and Gram classification. Green, fast method.	Dataset size not stated. Needs NIR accessory. Limited species.
Thomsen *et al.* (40), 2022	Spectral Transformer for Raman hyperspectral images. Dataset: 15 classes (6 MR-MS species).	96% accuracy (15 classes), 95.6% (MR-MS). Outperformed CNNs.	Dataset unspecified. Requires hyperspectral imaging. Limited to phenotypic classes.
Lu *et al.* (41), 2023	Raman + ML for species ID and resistance detection.	90.73% species-level, 99.92% resistance accuracy.	Dataset not provided. Single-cell analysis adds complexity.

	Dataset: 12 species, A. baumannii strains.		
Safir *et al.* (42), 2023	Acoustic bioprinting + SERS + ML. Dataset: S. epidermidis, E. coli, blood mixtures.	≥99% accuracy (pure), ≥87% (mixed). High enhancement (1500×).	Small dataset. Complex setup. Limited to select mixtures.
K. Kukula *et al.* (43), 2021	4-layer CNN for Raman spectra. Dataset: 30 bacterial classes.	86% accuracy. Reduced model complexity. Near real-time.	Moderate accuracy. Needs large spectral datasets.
L. Deng *et al.* (44), 2022	Deep NN with multi-receptive fields for Raman spectra. Dataset: Not specified.	Higher accuracy than prior methods. Visualization for interpretability.	Dataset missing. Limited clinical testing. Expert input needed.
Yichen Liu *et al.* (45), 2024	Wavelet packet + Gramian angular field + DL. Dataset: 2 and 30 isolates.	99.64% (2 isolates), 90.55% (30 isolates). Training time reduced 90%.	Lower accuracy on 30 classes. Dataset size unclear. Sensitive to noise.

The Spectroscopy-based AI demonstrates strong potential for bacterial identification, with studies reporting accuracies between 85.8% and 100%. Techniques such as NIR spectroscopy with PCA/HCA/KNN, SERS with acoustic bioprinting, and wavelet-based deep learning have achieved near-perfect results, including rapid detection of antibiotic resistance. Advanced models like spectral transformers and multi-receptive field DNNs further improve performance and interpretability. However, progress is limited by small or unspecified datasets, narrow species coverage, and moderate accuracy in multi-class tasks. The High equipment costs, spectral variability, and sensitivity to noise also hinder practical use. While methods like data augmentation and noise superposition improve robustness, broader adoption will require standardized datasets and cost-effective, accessible systems.

Frequent Techniques Used

Spectroscopy-based bacterial identification applies AI to analyze molecular signatures with high precision. Principal Component Analysis (PCA) and wavelet transforms reduce spectral dimensionality, while classifiers such as Support Vector Machines (SVM), Random Forest (RF), and Logistic Regression enable efficient classification. Convolutional and Deep Neural Networks (CNNs, DNNs) handle complex Raman and FTIR spectra, with advanced variants like spectral transformers and multi-receptive field models improving feature extraction. Additional approaches include K-Nearest Neighbor (KNN), Hierarchical Cluster Analysis (HCA), and image processing for susceptibility testing. Enhancements such as Surface-Enhanced Raman Spectroscopy (SERS) and acoustic bioprinting boost signal intensity and detection accuracy. Together, these methods enable rapid and reliable bacterial classification.

Applications

AI-enhanced spectroscopy supports diverse fields of microbiology. In clinical diagnostics, it enables rapid species identification and antibiotic resistance profiling, including drug-resistant strains like Acinetobacter baumannii. Automated antimicrobial susceptibility testing accelerates therapy decisions. Food safety benefits from detecting pathogens such as E. coli and Salmonella, while environmental monitoring identifies contaminants in water and soil. Applications also include Gram-positive/negative classification, distinguishing MR and MS strains for infection control, and single-cell pathogen detection in hospital settings. Beyond healthcare, spectroscopy-based AI aids public health surveillance and biomedical research by analyzing biomolecular resistance patterns.

Advantages

These techniques deliver high accuracies (85.8–100%), ensuring dependable results across applications. Non-destructive analysis preserves samples, while minimal preparation reduces time and labor. Rapid, often real-time detection supports timely interventions. High specificity allows differentiation of closely related strains, and SERS provides signal enhancement up to 1500×, boosting sensitivity. Deep learning extends capabilities to complex datasets while NIR spectroscopy offers sustainable, “green” testing options. Single-cell resolution and visualization of spectral features further support clinical decision-making. Robustness to noise, as shown in recent studies, strengthens reliability in real-world use.

Challenges

Despite progress, several obstacles remain. Many studies rely on small or unspecified datasets, limiting robustness and generalizability. Narrow species coverage further restricts applicability. Spectral variability from environmental or chemical factors reduces consistency. The need for specialized equipment, such as hyperspectral Raman or FO-FTIR, adds cost and complexity. Performance in multi-class tasks can be moderate, with accuracies of 85–86% in some studies. Computational demands of advanced AI models hinder deployment in resource-limited settings. Finally, the absence of standardized spectral databases and reliance on expert infrastructure remain barriers to large-scale adoption. Addressing these challenges requires larger, well-annotated datasets, cost-effective instrumentation, and simplified workflows for broader real-world implementation.

Hyperspectral Imaging and AI Techniques

Hyperspectral imaging (HSI) combined with AI enables precise bacterial identification by capturing both spectral and spatial data across a wide wavelength range, outperforming traditional culture-based methods in speed and specificity. When paired with models such as LSTM and deep neural networks, HSI can rapidly process complex datasets, supporting single-cell analysis and real-time applications. Recent studies demonstrate its effectiveness in detecting foodborne pathogens and analyzing mixed bacterial samples (13, 16). However, practical deployment is hindered by high equipment costs and limited dataset availability.

Table 4: Hyperspectral Imaging and AI Techniques

Author Details	Proposed Work and Dataset	Achievements	Limitations
Xinggong Liang *et al.* (46), 2023	AI classification of bacterial infections from pathology images at patch and whole slide levels. Dataset: Pathology images (size not specified).	AUC >0.950 in all phases. High accuracy and robustness.	Dataset size not disclosed. Low specificity for bacterial subtypes.
Hikaru Tago *et al.* (47), 2022	Line image sensor for colony fingerprinting + ML. Dataset: 15 species from 9 genera (size not provided).	96% accuracy in 10 hours. Petri dish scanned in 22 seconds. Faster than 24-h MS.	Limited to 15 species. Required 10-hour incubation. Dataset size not specified.
Zhu *et al.* (48), 2023	Hyperspectral Transmission Microscopic Imaging (HTMI) + PCA-SVM for single-cell classification. Dataset: Five bacterial species at low concentrations.	93.6% accuracy. Achieved 2.19 µm spatial and <1 nm spectral resolution.	Small dataset. Only five species. High computational demands.
Rui Kang *et al.* (49), 2022	HMI + LSTM for classifying five foodborne pathogens using ROIs (whole-cell, boundary, center).	92.9% accuracy (center ROI). Outperformed PCA-based methods (66–85%).	Limited to five pathogens. Dataset size moderate. Requires high-quality HMI.
Zhu *et al.* (50), 2024	Dual-mode HSI with MB-Net (deep neural network) to predict proportions of mixed bacteria. Dataset: Four mixed pathogenic species.	R^2 = 0.96, RMSE = 0.03. First method for simultaneous detection of 4 mixed bacteria.	Only four species. MB-Net is computationally intensive. Dataset size not stated.
Park *et al.* (51), 2023	AGR2U-Net + ellipse fitting for single-cell segmentation in FPI-	94.1% mIoU, 97.4% ellipse fitting accuracy. Robust to blurriness.	Limited to four species. High cost of FPI-HMI.

	HMI. Dataset: E. coli, Listeria, Salmonella, Staphylococcus.		Dataset moderately sized.

High potential for bacterial identification has been demonstrated by HSI combined with AI, achieving accuracies of 92.9–96% across diverse applications. An AUC above 0.95 has been reported for the classification of bacterial infections from pathology images. An accuracy of 96% has been achieved in the identification of 15 bacterial species within 10 hours, outperforming the 24-hour requirement of mass spectrometry. An accuracy of 93.6% has been obtained for single-bacterium classification using hyperspectral transmission microscopic imaging (HTMI). Five foodborne pathogens have been classified with 92.9% accuracy using LSTM networks, surpassing the performance of traditional PCA-based methods. Four mixed bacteria have been simultaneously detected with an R^2 of 0.96 using a custom DNN-based MB-Net. Mean intersection over union (IoU) of 94.1% and ellipse fitting accuracy of 97.4% have been achieved in single-cell segmentation, demonstrating robustness to image blurriness.

Limitations

Many studies are conducted using small or unspecified datasets, which limits generalizability. Research is often focused on a few bacterial species, reducing applicability to broader microbial diversity. High costs of HSI systems and the computational demands of AI models pose significant barriers. Implementation is complicated by dependence on high-quality data and incubation time. Despite advances in addressing image blurriness and mixed samples, the development of larger standardized datasets, cost-effective hardware, and efficient algorithms is required for practical clinical and industrial deployment.

Frequent Techniques Used

- Imaging Systems: HTMI for high spatial and spectral resolution; Fabry-Perot Interferometer (FPI) HSI for enhanced spectral resolution; line image sensors for rapid colony fingerprinting.
- AI and Machine Learning: PCA with SVM for spectral classification; LSTM networks for ROI spectral data; DNNs with spectral feature fusion for mixed bacteria (MB-Net 64); U-Net, ResU-Net, AGR2U-Net for single-cell segmentation.
- Data Processing: Deblurring, padding, and ellipse fitting enhance single-cell identification. AI also processes whole slide images and patch-level pathology data.

Applications

- Clinical Diagnostics: Identification of bacterial infections from pathology images.
- Food Safety: Detection of foodborne pathogens.
- Single-Bacterium and Mixed Sample Analysis: Precise pathogen identification and quantification in complex samples.
- Environmental Monitoring: Detecting microbial contaminants in food.
- Biomedical Research: Analyzing bacterial morphology and species diversity using spectral ROIs.
- Industrial and Workflow Optimization: Rapid colony fingerprinting and automated single-cell segmentation streamline testing and diagnostics.

Advantages

- High accuracy (92.9–96%) and AUC >0.95 for reliable classification for several studies.
- Single-cell resolution and high spatial (<2.19 μm) and spectral (<1 nm) resolution.
- Rapid diagnostics (e.g., 22 seconds per Petri dish).
- Robustness to blurriness and mixed bacteria detection ($R^2 = 0.96$).
- AI models outperform traditional methods (7–26% improvement over PCA-based classifiers).
- Non-invasive imaging preserves samples and supports real-time processing for clinical workflows.
- Scalable systems suitable for industrial applications.

Sensor Array and Machine Learning for Bacterial Detection

Sensor array technology combined with machine learning (ML) provides a rapid, cost-effective alternative to traditional microbiological methods, which are often time-consuming, labor-intensive, and require specialized laboratory setups. Fluorescence-based sensor arrays, using elements like carbon quantum dots (CQDs) or two-dimensional nanomaterials, generate distinct response patterns for different bacterial species. When paired with ML algorithms, these arrays analyze complex fluorescence fingerprints with high accuracy, enabling high-throughput identification. Recent studies employ cross-reactive receptors, fluorescence quenching, and advanced ML models to detect multiple bacterial species, including pathogens and drug-resistant strains. Applications span food safety, clinical diagnostics, and environmental monitoring. Despite their advantages, issues such as sensor cross-reactivity and limited species diversity remain, highlighting areas for further optimization. Table 5 summarizes the techniques, achievements, and limitations, providing a comprehensive overview of this emerging approach.

Table 5: Sensor Array and Machine Learning for Bacterial Detection

Author Details	Proposed Work and Dataset	Achievements	Limitations
Yi Wang *et al.* (52), 2023	Developed a fluorescence sensor array with 2D nanoparticles and ssDNA for identifying eight bacteria in milk. Dataset: Eight pathogenic and spoilage bacteria (size not specified).	Achieved 93.8% accuracy (30-min incubation), 98.4% with multilayer perceptron (120-min). Low-cost alternative to ELISA.	Dataset size not specified. Limited to eight species. Incubation time required.
Laibao Zheng *et al.* (53), 2022	Used fluorescence sensor array with carbon dots (boronic acid, polymixin, vancomycin) and LDA. Dataset: Six bacterial species (size not specified).	Demonstrated effective discrimination of six bacteria using fluorescence patterns. Simple and rapid method.	Dataset size not provided. Limited to six species. Cross-reactivity of receptors not quantified.
Li, Z (54), 2023	Developed six-sensing array with 2D nanomaterials and ssDNA for microbial identification. Dataset: Eight microorganisms, n=288 samples.	Achieved 97.9% accuracy across eight species, including drug-resistant strains. Rapid detection at low concentrations (10^2–10^8 CFU/mL).	Small dataset (n=288). Limited to eight species. Complex sensor fabrication.
Wang *et al.* (55), 2024	Developed paper-based fluorescence sensor array with antibiotic-modified CQDs and smartphone imaging. Dataset: Five bacterial strains (10^3–10^7 CFU/mL).	Differentiated five strains with high accuracy. Cost-effective, portable platform validated with blind samples.	Limited to five species. Moderate dataset size. Smartphone imaging quality variability.

The significant potential of sensor array technology combined with machine learning for bacterial detection has been highlighted, with high accuracies of 93.8–98.4% reported and practical applications demonstrated. An accuracy of 98.4% has been achieved in identifying eight bacteria in milk using a fluorescence sensor array with a multilayer perceptron, providing a low-cost alternative to ELISA. An accuracy of 97.9% has been obtained across eight microorganisms, including drug-resistant strains, using a six-sensing array capable of detecting low concentrations (10^2–10^8 CFU/mL). Six bacteria have been effectively discriminated using carbon dot-based sensors and linear discriminant analysis, emphasizing simplicity and speed. A portable paper-based platform with antibiotic-modified CQDs has been developed to accurately differentiate five bacterial strains, with validation performed using blind samples and smartphone imaging. These results demonstrate the precision, speed, portability, and cost-effectiveness of sensor arrays. However, limitations remain. Small or unspecified dataset sizes are still used, receptor cross-reactivity affects specificity, incubation times of 30–120 minutes are required, sensor fabrication is complex, and variability in smartphone imaging quality impacts performance. While some studies have addressed portability and low-concentration detection, the development of larger standardized datasets, improved receptor specificity, and simpler fabrication processes is required for widespread adoption in clinical, food safety, and environmental applications.

Frequent Techniques Used

Advanced approaches for bacterial identification are employed by sensor array methods combined with machine learning. Fluorescence-based arrays are constructed using single-stranded DNA (ssDNA) quenched by two-dimensional nanomaterials or carbon quantum dots (CQDs) functionalized with receptors such as boronic acid, polymyxin, vancomycin, or antibiotics. Bacterial surface interactions are monitored through aggregation-induced fluorescence quenching. Classification is performed using Linear Discriminant Analysis (LDA), multilayer perceptrons, and artificial neural networks, with SVM and K-Nearest Neighbors applied as baselines. Fluorescent signals collected at specific wavelengths (e.g., 520 nm) are used to generate microbial fingerprints. Portable detection is enabled by paper-based platforms with inkjet-printed CQDs, while smartphone imaging facilitates on-site analysis. Data preprocessing, including signal normalization, is applied to improve model performance. These techniques allow rapid and high-accuracy bacterial identification.

Applications

Wide-ranging applications are supported by sensor arrays combined with machine learning. Food safety is enhanced through the detection of pathogenic and spoilage bacteria in milk. Clinical diagnostics are facilitated for pathogens such as Escherichia coli, Staphylococcus aureus, and Klebsiella pneumoniae, including drug-resistant strains like MRSA. Environmental monitoring is performed to detect bacteria in water and soil, while portable paper-based platforms with smartphone integration enable point-of-care testing. Rapid microbial quality control benefits industrial microbiology, and microbial taxonomy and antibiotic interactions are studied in biomedical research using these methods. Real-world performance is validated through blind sample testing.

Advantages

These techniques achieve high accuracy (93.8–98.4%) and rapid detection within 30–120 minutes. They detect low concentrations (10^2–10^8 CFU/mL) and are cost-effective, especially on paper-based platforms. Portability with smartphone integration enables on-site testing, while non-specific cross-reactive receptors simplify sensor design. Minimal sample preparation, high specificity for drug-resistant strains, scalability for high-throughput applications, ease of fabrication via inkjet printing, and robustness to blind samples enhance practical utility.

Challenges

Key limitations include small or unspecified datasets restricting generalizability, focus on a limited number of species, receptor cross-reactivity reducing specificity, incubation times of 30–120 minutes, complex sensor fabrication, and variability in smartphone imaging quality. Dependence on stable fluorescence signals, lack of standardized datasets, and limited validation on mixed or complex samples further constrain implementation. Addressing these challenges requires larger datasets, optimized receptor specificity, and streamlined fabrication for broader adoption.

Other Advanced Imaging and AI Techniques

Using spectroscopy with AI is a modern way to identify bacteria. It allows fast and non-destructive testing. This helps with both classification and spotting antibiotic resistance. Spectroscopy methods—such as Raman, FTIR, and excitation-emission scans—pick up unique molecular patterns in bacteria. They are now used in clinical labs, food safety checks, and environmental studies. This section reviews 14 studies that use AI and spectroscopy to identify bacteria. These studies utilize advanced algorithms, including principal component analysis (PCA), convolutional neural networks (CNNs), and spectral transformers, to analyze spectral data. Applications range from pathogen detection to antimicrobial susceptibility testing, with significant advantages in speed and precision. However, challenges such as spectral variability, dataset limitations, and equipment costs persist. The details of the frequent techniques, applications, advantages, and limitations are discussed in the following Table 6, providing a comprehensive overview of this innovative field.

Table 6: Other Advanced Imaging and AI Techniques

Author Details	Proposed Work and Dataset	Achievements	Limitations
Ma L. *et al.* (56), 2023	Used YOLOv4 with phase-contrast microscopy for *E. coli* detection. Dataset: *E. coli* and seven foodborne bacteria, romaine lettuce samples.	Achieved 94% precision, R^2 = 0.995 for quantification, <10% false-negative rate. Rapid 3-h detection.	Limited to eight species. Requires 3-h cultivation. Dataset size not specified.
Hiraoka M. *et al.* (57), 2018	Developed image processing system for filamentous bacteria in wastewater. Dataset: Activated sludge samples.	Automated bulking control. Supported operator identification. Monitored control effects.	Dataset size not specified. Limited to filamentous bacteria. Requires operator interaction.
Demirel M. *et al.* (58), 2022	Used iterative Bayesian model for FLIM bacterial detection. Dataset: Synthetic and real FLIM images.	Improved F1 score by 16.85% over existing methods. Outperformed on real images.	Dataset size not specified. Computationally intensive. Limited to FLIM imaging.
Thomas Mortier *et al.* (59), 2021	Benchmarked ML methods for MALDI-TOF spectra. Dataset: 100,000 spectra, >1000 species.	Acceptable identification rates for novel replicates, strains, species. Used neural networks with Monte Carlo dropout.	Lower accuracy for novel species. Taxonomic information poorly preserved. Large dataset required.
Harris *et al.* (60), 2024	Reviewed ML for UTI diagnostics and antibiotic resistance prediction. Dataset: Literature-based.	Highlighted rapid diagnostics and reduced antibiotic use. Improved clinical workflows.	No experimental results. Relies on secondary data. Broad scope lacks specificity.
He *et al.* (61), 2023	Used SIM with ML for bacterial identification. Dataset: *E. coli*, *M. smegmatis*, *P. aeruginosa* images.	Achieved 98% classification accuracy. Rapid morphological analysis.	Limited to three species. Small dataset. Requires high-resolution SIM imaging.
Ragi S. *et al.* (62), 2023	Used DCNN and Mask R-CNN for SEM image segmentation of DA-G20 cells. Dataset: SEM images.	70–227× faster than manual methods. Accurate geometric property extraction.	Dataset size not specified. Limited to DA-G20 cells. Requires SEM expertise.
Ding Y *et al.* (63), 2024	Developed paper-based fluorogenic probe with smartphone AI for β-lactamase detection. Dataset: Bacterial samples, mice.	Detected β-lactamase in 20 s, 0.13 nmol/L detection limit. Calibrated for complex samples.	Dataset size not specified. Limited to β-lactamase. Smartphone variability.
Ahmad N. *et al.* (64), 2019	Used 3-layer neural network for Peptococcaceae identification. Dataset: Bergey's manual data.	Rapid identification vs. manual methods. High accuracy for facultative anaerobes.	Dataset size not specified. Limited to Peptococcaceae. Relies on manual data.
Kim G. *et al.* (65), 2022	Used CNN for 3D refractive index image classification. Dataset: Not specified.	Rapid and accurate species identification.	Dataset details absent. Limited species scope.

		Simplified microbial detection.	Requires 3D imaging setup.
Zhang S. *et al.* (66), 2023	Used DL for impedance-based analysis of three bacteria. Dataset: *EPEC*, *S. enteritidis*, *V. parahaemolyticus*.	Achieved 100% accuracy. Suitable for point-of-care testing.	Small dataset. Limited to three species. Requires impedance system.
Demirel M. *et al.* (67), 2023	Generated synthetic bacteria in OEM images using 3D U-Net. Dataset: Synthetic and real OEM images.	Improved correlation by 3.86% over baseline. Enhanced detection with DLNet.	Limited dataset size. Synthetic data may not fully represent real variability.
Nirmala Bai L. *et al.* (68), 2024	Used EfficientNetB1/B2 for chest infection diagnosis from X-rays. Dataset: Public and hospital X-ray images.	Superior accuracy vs. other transfer learning models. Effective for infections.	Dataset size not specified. Limited to chest infections. Requires X-ray imaging.
Yang Zhang *et al.* (69), 2023	Reviewed DL for microbial image analysis. Dataset: Literature-based.	Highlighted potential for viruses, bacteria, fungi, parasites. Guided future research.	No experimental results. Relies on secondary data. Broad scope lacks depth.
Maaskant *et al.* (70), 2024	Used DL/ML for fecal smear bacterial prediction. Dataset: Rhesus macaque fecal images.	Predicted 16 genera (AUC >0.7), butyrate producers (AUC 0.75). Robust to noise.	Dataset size not specified. Limited to fecal bacteria. Requires metagenomic data.

The transformative impact of advanced imaging and AI techniques on bacterial detection is highlighted, with accuracies of 94–100% achieved and innovative applications demonstrated across various domains. A precision of 94% was obtained in detecting E. coli using YOLOv4 with phase-contrast microscopy, enabling rapid food safety testing in just 3 hours. An accuracy of 98% was achieved in classifying three bacterial species using SIM and machine learning, demonstrating the potential of morphological analysis. An accuracy of 100% was attained in identifying three pathogens via impedance-based deep learning, suitable for point-of-care testing. F1 scores were improved by 16.85% using a Bayesian model for FLIM, outperforming traditional methods. Machine learning was benchmarked on a dataset of 100,000 spectra, achieving acceptable rates for novel species identification. SEM analysis was accelerated 70–227 times using DCNN and Mask R-CNN, enhancing biofilm research. However, limitations remain, including the use of small or unspecified datasets restricting generalizability, a focus on few species limiting microbial diversity, costly and complex imaging systems such as SIM, FLIM, and impedance setups, and the computational demands of models like 3D U-Net and EfficientNet. The lack of experimental data in review studies reduces specificity. While synthetic data and portable probes address some challenges, standardized datasets, cost-effective systems, and broader species validation remain critical for adoption in clinical, industrial, and environmental settings.

Frequent Techniques Used

Diverse AI approaches are employed in advanced imaging methods. Bacteria are detected in phase-contrast microscopy using YOLOv4, FLIM data are analyzed with Bayesian models and Metropolis-Hastings sampling, and SIM and 3D refractive index images are classified with CNNs. SEM images are segmented using DCNNs and Mask R-CNNs, while X-ray images for infection detection are analyzed with EfficientNet models. Novel species in MALDI-TOF spectra are identified using neural networks with Monte Carlo dropout, and bacteria in synthetic optical endomicroscopy images are detected with 3D U-Net models. Bacterial subtypes are classified using impedance-based analysis combined with deep learning, and β-lactamase is detected with paper-based fluorogenic probes supported by smartphone AI. Additional methods include hierarchical classification of spectral datasets, geometric analysis with moment invariants, feed-forward backpropagation, and explainability analysis to improve model transparency.

Applications

These techniques have broad applications. Food safety is enhanced through rapid pathogen detection, such as E. coli in romaine lettuce, while clinical diagnostics identify pathogens including E. coli, Salmonella, and Vibrio parahaemolyticus and support urinary tract infection diagnosis and antibiotic resistance prediction. Environmental monitoring detects filamentous bacteria in wastewater, and biomedical research investigates biofilm phenotypes and β-lactamase activity. Industrial microbiology benefits from rapid species identification for quality control, point-of-care testing is enabled by portable probes and impedance systems, and novel species identification advances microbial taxonomy. Real-time pathogen monitoring and fecal bacterial group prediction support infection control and gut health studies.

Advantages

High accuracy (94–100%) ensures reliable identification, with rapid detection times and low false-negative rates (<10%) enhancing diagnostic reliability. Non-invasive imaging preserves samples, while high specificity distinguishes bacterial subtypes and resistant strains. Automation reduces manual labor, and scalability handles large datasets, such as 100,000 spectra. Portability via smartphone integration supports point-of-care testing, and robustness to noise improves practical utility. High quantification accuracy (R^2 = 0.995), explainability analysis, versatility across imaging modalities, and significant speed improvements up to 227× faster than manual methods further demonstrate their transformative potential.

Table 7 presents the distribution of studies across different research topics, highlighting the prominence of various AI-driven approaches for bacterial detection and classification. The majority of 31 studies focus on microscopic image-based classification using machine learning and deep learning, reflecting its dominant role in this field. Spectroscopy-based identification using AI has also gained considerable attention, with 14 studies, while other advanced imaging and AI techniques account for 15 studies, showcasing the exploration of diverse methods beyond conventional imaging. In comparison, hyperspectral imaging and AI techniques have been applied in 6 studies, and sensor array-based approaches with machine learning have been examined in 4 studies. This distribution indicates that while image-based and spectroscopy-driven methods remain central, emerging techniques are gradually expanding the research landscape.

Table 7: Topic Wise Studies Count

Topic	Study Count
Microscopic Image-Based Classification Using Machine Learning and Deep Learning	31
Spectroscopy-Based Identification Using AI	14
Hyperspectral Imaging and AI Techniques	6
Sensor Array and Machine Learning for Bacterial Detection	4
Other Advanced Imaging and AI Techniques	15

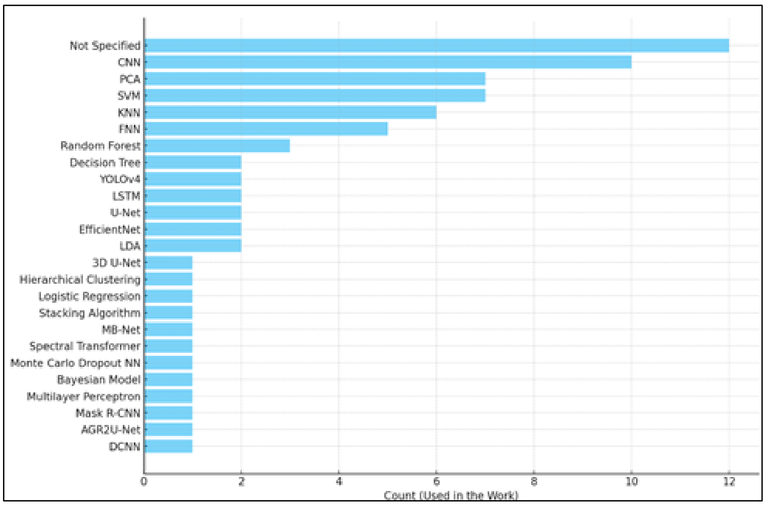

Figure 4: Dataset Used in this Study

Figure 4 shows the usage frequency of different models across studies. Most works either did not specify the model-5 or used CNNs-4, followed by PCA-7 and SVMs-6. KNN, FNN, and random forests appeared less often, while advanced models like U-Net, YOLOv4, EfficientNet, and LSTM were used in limited cases.

Limitation and Challenges

AI-based bacterial identification techniques face significant challenges that hinder their widespread adoption. Small or unspecified dataset sizes limit model generalizability. Many studies rely on datasets with fewer than 1000 samples, such as 22 sputum smear images or 72 spectra. Others omit dataset details entirely, hindering reproducibility. A narrow focus on specific bacterial species further restricts applicability. Research often targets three to eight species, such as E. coli or Staphylococcus aureus, neglecting the diverse microbial populations. This specificity reduces utility in real-world settings with complex microbiomes. Lack of standardized, large-scale datasets across techniques exacerbates these issues, complicating model training and validation. High computational complexity poses a significant barrier to progress. Advanced models, including YOLOv4, Swin Transformer, and 3D U-Net, demand substantial processing power. Deep neural networks and spectral transformers require extensive resources, which limit their deployment in resource-constrained environments, such as rural clinics. Specialized equipment adds to accessibility challenges. High-resolution microscopy, hyperspectral sensors, and fluorescence lifetime imaging systems are costly.

Data variability undermines model reliability. Inconsistent staining, image blurriness, or environmental factors affect imaging quality. Spectral variability from dye interactions or noise impacts spectroscopy results. Fluorescence signal stability in sensor arrays depends on controlled conditions. Robust preprocessing is often lacking, which can reduce performance in suboptimal settings. Validation on mixed or complex samples is limited. Many studies focus on single-species or low-concentration samples, ignoring real-world diversity. Synthetic data, while innovative, may not fully capture natural variability, requiring further testing. Moderate accuracies in multi-class tasks, ranging from 85.8% to 86%, highlight the difficulties in scaling to diverse scenarios.

Critical Appraisal and Risk of Bias Assessment

To ensure the reliability of findings in this systematic review, a critical appraisal of the 70 included studies was conducted to assess the risk of bias. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool, adapted for use in AI-based diagnostic studies, was employed. This tool evaluates four domains: Patient/Sample Selection, Index Test, Reference Standard, and Flow and Timing. Each domain was assessed for risk of bias and concerns regarding applicability. Two independent reviewers (N.A.J. and G.P.G.) performed the appraisal. Discrepancies were resolved through discussion or consultation with a senior reviewer (C.J.). The assessment aimed to identify methodological weaknesses that could affect the validity of reported outcomes, such as classification accuracies in bacterial identification.

Methods of Critical Appraisal

The QUADAS-2 tool was tailored to address AI-based bacterial identification studies. The domains were defined as follows:

- **Patient/Sample Selection**: Evaluated whether datasets were representative of real-world bacterial populations. Studies with small, non-diverse datasets (e.g., 3–8 species) or unclear sampling methods were rated as high risk. Applicability concerns arose if datasets did not reflect clinical, food safety, or environmental contexts.
- **Index Test**: Assessed the clarity and reproducibility of AI algorithms (e.g., machine learning, deep learning) and imaging/sensing techniques (e.g., spectroscopy, hyperspectral imaging). Studies lacking detailed model descriptions or validation methods were rated as high risk.
- **Reference Standard**: Examined the accuracy of ground truth labels (e.g., confirmed bacterial species via culture or molecular methods). Studies with unclear or unverified reference standards were rated as high risk.
- **Flow and Timing**: Evaluated consistency in applying AI and sensing methods across samples. Studies with incomplete reporting of testing protocols or inconsistent application were rated as high risk.

  Each study was rated as low, high, or unclear risk of bias for each domain. Applicability concerns were noted separately to assess relevance to the review's objectives. Results were summarized in a table (Table 7) and narratively synthesized to highlight trends and implications.

Results of Risk of Bias Assessment

Of the 70 studies, 40% (28 studies) were rated as low risk across all domains, indicating robust methodology. However, 50% (35 studies) had high or unclear risk in at least one domain, primarily due to:

- **Sample Selection**: 25 studies used small datasets (e.g., <100 samples) or focused on few bacterial species, limiting generalizability. Only 15 studies included diverse, multi-species datasets relevant to clinical or environmental applications.
- **Index Test**: 20 studies provided insufficient details on AI model parameters (e.g., hyper-parameters, training protocols) or lacked external validation, increasing bias risk.
- **Reference Standard**: 10 studies had unclear reference standards, relying on unverified labels or proprietary datasets, reducing reliability.
- **Flow and Timing**: 15 studies incompletely reported testing protocols, such as inconsistent imaging or sensor application, raising concerns about reproducibility.

Applicability concerns were noted in 30 studies, particularly those with datasets not aligned with real-world applications (e.g., lab-based samples versus clinical isolates).

Implications

The risk of bias assessment revealed strengths and weaknesses. Studies with low risk provided reliable evidence of AI's high accuracy (85.8%–99%) in bacterial identification. However, high or unclear risk in sample selection and index test domains suggests caution in interpreting results from studies with small or poorly described datasets. These limitations may overestimate performance in real-world settings. The assessment informed the review's synthesis, prioritizing findings from low-risk studies. Future research should focus on larger, diverse datasets and transparent reporting of AI methods to reduce bias.

Future Direction and the Approaches

Advancing AI-based bacterial detection requires addressing key limitations in datasets, computational demands, hardware accessibility, and real-world applicability. Small or unspecified datasets limit model generalizability, making the development of large, standardized, and open-source datasets essential. These datasets should encompass diverse bacterial species, mixed samples, and real-world conditions, with metadata such as staining protocols or imaging parameters to enhance reproducibility. Collaborative platforms, global microbial databases, and public-private partnerships can facilitate data sharing and funding. Synthetic data generation using advanced models, such as GANs, validated against real data, and crowdsourced expert annotations can further improve dataset quality.

Reducing computational complexity is crucial for practical deployment, especially in resource-limited settings. Lightweight AI architectures, like MobileNet or depth-wise separable CNNs, can maintain accuracy while reducing processing requirements. Techniques such as quantization, pruning, edge computing, and federated learning allow on-device processing and preserve data privacy, particularly in clinical contexts. Open-source frameworks, such as TensorFlow Lite, can accelerate adoption of efficient models. Cost-effective and portable imaging systems are essential for broader adoption. High-cost equipment, including hyperspectral sensors or FLIM, limits access, whereas smartphone-based microscopy, paper-based fluorescence sensor arrays, portable Raman or FTIR devices, and 3D-printed imaging components offer affordable alternatives. Multispectral imaging can balance cost and performance while maintaining sufficient resolution for reliable detection. Enhancing model robustness to data variability is critical, using techniques such as adaptive normalization, domain adaptation, transfer learning, ensemble learning, noise-robust algorithms, and real-time data augmentation.

Expanding validation to mixed and complex microbial samples is a priority. Most studies focus on single-species or low-concentration samples, but real-world scenarios require hierarchical classification, multi-task learning, and longitudinal studies to assess performance across diverse environments. Explainable AI methods, such as attention maps, increase trust and interpretability. Optimizing receptor specificity for sensor arrays, using bioinformatics, high-throughput screening, and hybrid imaging-spectroscopy models, can reduce cross-reactivity and improve accuracy. Eliminating cultivation delays is another critical goal. Even short incubation times (3–10 hours) impede rapid diagnostics, highlighting the need for culture-free techniques, including impedance-based analysis, single-cell spectroscopy, real-time AI-integrated imaging, and microfluidic isolation. Automation of sample preparation through robotics can further accelerate processing, enabling near-instantaneous bacterial identification for urgent clinical settings.

Finally, promoting interdisciplinary collaboration and accessible expertise is vital. Training programs, online courses, virtual workshops, and open-source AI tools can empower microbiologists, clinicians, and engineers. Interdisciplinary teams and global networks can standardize protocols, share best practices, and develop user-friendly interfaces, ensuring that advanced AI-based techniques are widely adopted in clinical diagnostics, food safety, and environmental monitoring.

Conclusion

AI-based bacterial identification techniques represent a groundbreaking advancement in microbiology, offering rapid, accurate, and automated alternatives to traditional methods. These approaches achieve exceptional accuracies, often ranging from 94% to 100%, across diverse applications. Advanced imaging modalities, such as phase-contrast microscopy and hyperspectral systems, capture detailed microbial signatures. Machine learning and deep learning models, such as YOLOv4 and convolutional neural networks, excel at classifying complex datasets. Spectroscopy techniques, including Raman and FTIR, provide non-destructive molecular analysis. Fluorescence-based sensor arrays enable portable detection. These methods significantly reduce identification times, from days to hours or minutes, compared to culture-based assays. They support critical applications in clinical diagnostics, food safety, and environmental monitoring. For instance, rapid pathogen detection aids timely treatment in hospitals. Automated systems enhance quality control in food industries. Point-of-care platforms, such as smartphone-integrated sensors, enable testing in resource-limited settings. The ability to detect antibiotic-resistant strains, quantify bacterial concentrations, and analyze mixed samples underscores the versatility of these techniques. Collectively, they address pressing global challenges, such as infectious disease management and antimicrobial resistance.

Despite their promise, significant challenges remain. Small or unspecified datasets limit model generalizability. Many studies focus on a few bacterial species, neglecting microbial diversity. The high computational demands of models like 3D U-Net or spectral transformers restrict deployment in low-resource environments. Costly equipment, such as hyperspectral sensors or FLIM systems, confines techniques to well-funded laboratories. Data variability, from staining inconsistencies to spectral noise, affects reliability. Cross-reactivity in sensor arrays reduces specificity. Incubation times, even reduced, delay results. Limited validation on mixed samples

hinders real-world applicability. The lack of standardized datasets and open-source resources hinders collaboration. Specialized expertise for operating complex systems poses barriers. These challenges underscore the disparity between research advancements and their practical application, particularly in underserved regions.
The path forward involves targeted innovations to overcome these barriers. Developing large, standardized, open-source datasets is essential. These should include diverse species and real-world conditions. Lightweight AI models, like MobileNet, can reduce computational demands. Affordable imaging systems, such as smartphone-based microscopy, can enhance accessibility. Robust preprocessing can address data variability. Validation on mixed samples can ensure the real-world utility of the model. Optimizing sensor specificity can improve detection precision. Culture-free techniques, like impedance-based analysis, can eliminate delays. Interdisciplinary collaboration, uniting microbiologists, engineers, and data scientists, can drive progress. Training programs can build expertise. Global consortia can standardize protocols. These efforts will bridge the gap between innovation and adoption, enabling AI techniques to transform bacterial identification.

Abbreviations

AI: Artificial Intelligence, CNN: Convolutional Neural Networks, CPN: Causal Probabilistic Network, DL: Deep Learning, FTIR: Fourier Transform Infrared, ML: Machine Learning, MRSA: Methicillin-Resistant Staphylococcus Aureus, PCR: Polymerase Chain Reaction.

Acknowledgement

None.

Author Contributions

T S Urmila: Data collection, screening of articles, Data extraction, J I Christy Eunaicy: Data collection, screening of articles, Data extraction, C Jayparatha: Methodology design, validation of findings, tabulation, technical validation, literature review, drafting specific sections, editing, preparation of figures and tables, Naveen Anandakumar J: Methodology design, validation of findings, tabulation, technical validation, literature review, drafting specific sections, editing, preparation of figures and tables, Govindaprabhu G B: Methodology design, validation of findings, tabulation, technical validation, literature review, drafting specific sections, editing, preparation of figures and tables, Mahalakshmi Priya R: Methodology design, validation of findings, tabulation, technical validation, literature review, drafting specific sections, editing, preparation of figures and tables.

Conflict of Interests

None.

Declaration of Artificial Intelligence (AI) Assistance

The authors declare no use of Artificial intelligence (AI) for the write-up of the manuscript.

Ethics Approval

All Ethics Followed and approved.

Funding

None.

References

1. Mohamed BA, Afify HM. Automated classification of bacterial images extracted from digital microscope via bag of words model. 9th Cairo International Biomedical Engineering Conference (CIBEC). IEEE Biomed Eng. 2018;1(2):86-89.
2. Kristensen K, Ward LM, Mogensen ML, Cichosz SL. Using image processing and automated classification models to classify microscopic gram stain images. Computer Methods and Programs in Biomedicine Update. 2023 Jan 1;3:100091.
3. Preetha V, Pandi Selvi P. Identification of bacteria using digital image processing. International Journal of Engineering Research in Computer Science and Engineering. 2018 Mar;5(3):254-8.
4. Wahid MF, Hasan MJ, Alom MS, Mahbub S. Performance analysis of machine learning techniques for microscopic bacteria image classification. Proc IEEE Int Conf Comput Commun Netw Technol. 2019;2(3):1-14.
5. Keren F, Keziah F, Gnanaraj FF, Vanitha L. Evaluation of automatic bacterial classification methods and approaches. Proc IEEE Int Conf Smart Struct Syst. 2023;1(1):1-6.
6. Kotwal S, Rani P, Arif T. Automated bacterial classifications using machine learning based computational techniques: architectures, challenges and open research issues. Arch Comput Methods Eng. 2022;29(3):2469-2490.
7. Sajedi H, Mohammadipanah F, Pashaei A. Image-processing based taxonomy analysis of bacterial macro-morphology using machine-learning models. Multimed Tools Appl. 2020;79(2):32711-32730.
8. Satyanarayana KV, Rao NT, Bhattacharyya D, Hu YC. Identifying the presence of bacteria on digital images by using asymmetric distribution with k-

means clustering algorithm. Multidim Syst Signal Process. 2022;33(3):301-326.

9. Rani P, Kotwal S, Manhas J, Sharma V, Sharma S. Machine learning and deep learning based computational approaches in automatic microorganisms image recognition: methodologies, challenges, and developments. Arch Comput Methods Eng. 2022;29(3):1801-1837.
10. Amitha IC, Rithul AK, Akash M, Amarnath T, Drishyap K. An AI enabled framework for bacteria detection from microscopic images for biomedical applications. Int J Novel Res Dev. 2024;9(7):688-698.
11. Singh A, Kumar A, Salman HM, Rawat N, Jain SK, Rao AT. Transfer learning approach on bacteria classification from microscopic images. Proc IEEE Int Conf Contemp Comput Informat. 2022;3(3):982-987.
12. Rani P, Kotwal S, Manhas J. Deep learning-based approach in automatic microscopic bacterial image classification. Proc Emerg Trends Expert Appl Secur. ICETEAS 2023. Lecture Notes in Networks and Systems, Singapore, Springer. 2023;4(2):1-12.
13. Kumar S, Mittal GS. Rapid detection of microorganisms using image processing parameters and neural network. Food Bioprocess Technol. 2010;3(5):741-751.
14. Hardo G, Noka M, Bakshi S. Synthetic micrographs of bacteria (SyMBac) allow accurate segmentation of bacterial cells using deep neural networks. BMC Biol. 2022;20(1):172-182.
15. Ramesh H, Elshinawy A, Ahmed A, Kassoumeh MA, Khan M, Mounsef J. BIOINTEL: Real-time bacteria identification using microscopy imaging. Proc IEEE Int Symp Biomed Imaging. IEEE Trans Med Imaging. 2024;1(2):1-4.
16. Yang M, Nurzynska K, Walts AE, Gertych A. A CNN-based active learning framework to identify mycobacteria in digitized Ziehl-Neelsen stained human tissues. Comput Med Imaging Graph. 2020;84(1):1-28.
17. Kotwal S, Rani P, Arif T, Manhas J, Sharma S. Machine learning and deep learning-based hybrid feature extraction and classification model using digital microscopic bacterial images. SN Comput Sci. 2023;4(3):687-700.
18. Sarker IA, Rahman MS, Hossain S, Ahmed F, Islam MN. Utilizing deep learning for microscopic image-based bacteria species identification. Proc IEEE Int Conf Comput Appl Syst. 2024;1(1):1-5.
19. Wahid MF, Hasan MJ, Alom MS. Deep convolutional neural network for microscopic bacteria image classification. Proc IEEE Int Conf Adv Electr Eng. 2019;3(2):866-869.
20. Ahmed T, Wahid MF, Hasan MJ. Combining deep convolutional neural network with support vector machine to classify microscopic bacteria images. Proc IEEE Int Conf Electr Comput Commun Eng. 2019;32(3):1-5.
21. Sunanda I, BK D, DV D, SK D, Sai Manusha GC. Classification of bacteria from agar plate using deep learning and image processing. Proc IEEE Int Conf Adv Data Eng Intell Comput Syst. 2024;32(3):1-6.
22. Visitsattaponge S, Bunkum M, Pintavirooj C, Paing MP. A deep learning model for bacterial classification using Big Transfer (BiT). IEEE Access. 2024;12(3):15609-15621.
23. Nasip ÖF, Zengin K. Deep learning-based bacteria classification. Proc IEEE Int Symp Multidiscip Stud Innov Technol. 2018;3(2):1-5.
24. Wahid MF, Ahmed T, Habib MA. Classification of microscopic images of bacteria using deep convolutional neural network. Proc IEEE Int Conf Electr Comput Eng. 2018;5(2):217-230.
25. Panicker RO, Kalmady KS, Rajan J, Sabu MK. Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng. 2018;38(3):691-699.
26. Andreini P, Bonechi S, Bianchini M, Mecocci A, Scarselli F. A deep learning approach to bacterial colony segmentation. Proc Artif Neural Netw Mach Learn. Lecture Notes in Computer Science, Springer. 2018;522-553. https://doi.org/10.1007/978-3-030-01424-7_51
27. Yang F, Zhong Y, Yang H, Wan Y, Hu Z, Peng S. Microbial colony detection based on deep learning. Appl Sci. 2023;13(19):10568-89.
28. Sengupta B, Alrubayan M, Kolla M, Wang Y, Mallet E, Torres A, Paniagua G, Krishnamurthi V. AI-based detection of optical microscopic images of Pseudomonas aeruginosa in planktonic and biofilm states. Information. 2025;16(4):309-321.
29. Iriya R, Braswell B, Mo M, Zhang F, Haydel SE, Wang S. Deep learning-based culture-free bacteria detection in urine using large-volume microscopy. Biosensors. 2024;14(2):89-97.
30. Mai DT, Ishibashi K. Small-scale depthwise separable convolutional neural networks for bacteria classification. Electronics. 2021;10(23):3005-3015.
31. Chin SY, Dong J, Hasikin K, Ngui R, Lai KW, Yeoh PSQ, Wu X. Bacterial image analysis using multi-task deep learning approaches for clinical microscopy. PeerJ Comput Sci. 2024;10(1):2180-2197.
32. Wan-dan Z, Ru-jin S, Cheng-wei W, Qian-xue L, Zhi-ping X. Raman spectroscopy classification of foodborne pathogenic bacteria based on PCA-stacking model. Proc Int Conf Intell Inform Biomed Sci. IEEE J Biomed Sci. 2019;1(1):304-307.
33. Ji SY, Jeong DH, Hassan M, Ilev IK. Signature infrared bacteria spectra analyzed by an advanced integrative computational approach developed for identifying bacteria similarity. IEEE J Sel Top Quantum Electron. 2019;25(1):720-729.
34. De Biasio M, McGunnigle G, Leitner R, Popp J, Rösch P, Balthasar D. Identification of single bacteria using micro-Raman spectroscopy. Proc Int Conf Sens Technol. IEEE Sens J. 2013;2(3):34-39.
35. Henry J, Endres JL, Sadykov MR, Bayles KW, Svechkarev D. Fast and accurate identification of pathogenic bacteria using excitation–emission spectroscopy and machine learning. Sens Diagn. 2024;3(8):1253-1262.
36. Barrera Patiño CP, Soares JM, Blanco KC, Bagnato VS. Machine learning in FTIR spectrum for the identification of antibiotic resistance: a demonstration with different species of microorganisms. Antibiotics. 2024;13(9):821-840.

37. Rahman MH, Sikder R, Tripathi M, Zahan M, Ye T, Gnimpieba ZE. Machine learning-assisted Raman spectroscopy and SERS for bacterial pathogen detection: clinical, food safety, and environmental applications. Chemosensors. 2024;12(7):140-160.
38. Gullu E, Bora S, Beynek B. Exploiting image processing and artificial intelligence techniques for the determination of antimicrobial susceptibility. Appl Sci. 2024;14(9):3950-3970.
39. Farias LR, Panero JS, Riss JSP, Correa APF, Vital MJS, Panero FS. Rapid and green classification method of bacteria using machine learning and NIR spectroscopy. Sensors. 2023;23(17):7336-7350.
40. Thomsen BL, Christensen JB, Rodenko O, Usenov I, Grønnemose RB, Andersen TE, Petersen E. Accurate and fast identification of minimally prepared bacteria phenotypes using Raman spectroscopy assisted by machine learning. Sci Rep. 2022;12(1):158-170.
41. Lu W, Li H, Qiu H, Wang L, Feng J, Fu YV. Identification of pathogens and detection of antibiotic susceptibility at single-cell resolution by Raman spectroscopy combined with machine learning. Front Microbiol. 2023;13(2):107-125.
42. Safir F, Vu N, Tadesse LF, Firouzi K, Banaei N, Jeffrey SS, Saleh AAE, Dionne JA. Combining acoustic bioprinting with AI-assisted Raman spectroscopy for high-throughput identification of bacteria in blood. Nano Lett. 2023;23(6):2065-2073.
43. Kukula K, Farmer D, Duran J, Majid N, Chatterley C, Jessing J, Li Y. Rapid detection of bacteria using Raman spectroscopy and deep learning. Proc IEEE Annu Comput Commun Workshop Conf. 2021;4(5):796-799.
44. Deng L, Zhong Y, Wang M, Zheng X, Zhang J. Scale-adaptive deep model for bacterial Raman spectra identification. IEEE J Biomed Health Inform. 2022;26(1):369-378.
45. Liu Y, Gao Y, Niu R, Zhang Z, Lu GW, Hu H, Liu T, Cheng Z. Rapid and accurate bacteria identification through deep-learning-based two-dimensional Raman spectroscopy. Anal Chim Acta. 2024;13(32):343376-91.
46. Liang X, Wang G, Zhu Z, Zhang W, Li Y, Luo J. Using pathology images and artificial intelligence to identify bacterial infections and their types. J Microbiol Methods. 2025;4(9):232-234.
47. Tago H, Maeda Y, Tanaka Y, Kohketsu H, Lim TK, Harada M, Yoshino T, Matsunaga T, Tanaka T. Line image sensor-based colony fingerprinting system for rapid pathogenic bacteria identification. Biosensors and Bioelectronics. 2024 Apr 1;249:116006.
48. Zhu H, Luo J, Liao J, He S. High-accuracy rapid identification and classification of mixed bacteria using hyperspectral transmission microscopic imaging and machine learning. Progress In Electromagnetics Research. 2023 Sep 1;178.https://doi.org/10.2528/PIER23082303
49. Kang R, Park B, Ouyang Q, Ren N. Rapid identification of foodborne bacteria with hyperspectral microscopic imaging and artificial intelligence classification algorithms. Food Control. 2021;130(12):108379.
50. Zhu H, Luo J, He S. Detecting multiple mixed bacteria using dual-mode hyperspectral imaging and deep neural networks. Appl Sci. 2024;14(4):1525-39.
51. Park B, Shin T, Kang R, Fong A, McDonogh B, Yoon SC. Automated segmentation of foodborne bacteria from chicken rinse with hyperspectral microscope imaging and deep learning methods. Comput Electron Agric. 2023;208:107802.
52. Wang Y, Feng Y, Xiao Z, Luo Y. Machine learning supported single-stranded DNA sensor array for multiple foodborne pathogenic and spoilage bacteria identification in milk. Food Chem. 2025;463(3):114-120.
53. Zheng L, Qi P, Zhang D. Identification of bacteria by a fluorescence sensor array based on three kinds of receptors functionalized carbon dots. Sens Actuators B Chem. 2019;286(2):206-213.
54. Li Z, Jiang Y, Tang S, Zhu L, Ma Y, Wang M, Li F. 2D nanomaterial sensing array using machine learning for differential profiling of pathogenic microbial taxonomic identification. Microchim Acta. 2022;189(2):273-285.
55. Wang F, Xiao M, Qi J, Zhang Y, Wang S. Paper-based fluorescence sensor array with functionalized carbon quantum dots for bacterial discrimination using a machine learning algorithm. Anal Bioanal Chem. 2024;416(2):3139-3148.
56. Ma L, Yi J, Wisuthiphaet N, Earles M, Nitin N. Accelerating the detection of bacteria in food using artificial intelligence and optical imaging. Applied and Environmental Microbiology. 2023 Jan 31;89(1):e01828-22.
57. Hiraoka M, Tsumura K, Enbutsu I, Yamamoto Y. Computer-based filamentous microorganism identification support system. Proc Int Workshop Artif Intell Ind Appl. IEEE Trans Ind Inform. 1988:283-288. https://doi.org/10.1109/AIIA.1988.13307
58. Demirel M, Mills B, Gaughan E, Dhaliwal K, Hopgood JR. Bayesian statistical analysis for bacterial detection in pulmonary endomicroscopic fluorescence lifetime imaging. IEEE Trans Image Process. 2024;33(1):1241-1256.
59. Mortier T, Wieme AD, Vandamme P, Waegeman W. Bacterial species identification using MALDI-TOF mass spectrometry and machine learning techniques: a large-scale benchmarking study. Comput Struct Biotechnol J. 2021;19(2):6157-6168.
60. Harris M. Machine learning and artificial intelligence for pathogen identification and antibiotic resistance detection: advancing diagnostics for urinary tract infections. BioMed. 2023;3(2):246-255.
61. He Y, Xu W, Zhi Y, Tyagi R, Hu Z, Cao G. Rapid bacteria identification using structured illumination microscopy and machine learning. J Innov Opt Health Sci. 2018;11(1):1850007-19.
62. Ragi S, Rahman MH, Duckworth J, Jawaharraj K, Chundi P, Gadhamshetty V. Artificial intelligence-driven image analysis of bacterial cells and biofilms. IEEE/ACM Trans Comput Biol Bioinform. 2023;20(1):174-184.
63. Ding Y, Chen J, Wu Q, Li X, Zhang Y, Wang L, Zhao W. Artificial intelligence-assisted point-of-care testing system for ultrafast and quantitative detection of

drug-resistant bacteria. SmartMat. 2024;5(1):214-230.
64. Ahmad N, Abdullah SRS, Anuar N, Husin H. Bacteria identification using artificial neural network: a case study of Peptococcaceae family identification. Proc Int Conf Electron Des. IEEE Conf Proc. 2008:1-6. https://doi.org/10.1109/ICED.2008.4786682
65. Kim G, Jo Y, Cho H, Choi G, Kim BS, Min HS, Park Y. Automated identification of bacteria using three-dimensional holographic imaging and convolutional neural network. Proc IEEE Photonics Conf. 2018;9(3):1-21.
66. Zhang S, Han Z, Feng Z, Sun M, Duan X. Deep learning assisted microfluidic impedance flow cytometry for label-free foodborne bacteria analysis and classification. Proc IEEE Annu Int Conf IEEE Eng Med Biol Soc. 2021;12(4):7087-7090.
67. Demirel M, Mills B, Gaughan E, Dhaliwal K, Hopgood JR. Bacteria detection in optical endomicroscopy images using synthetic images. Proc IEEE Annu Int Conf IEEE Eng Med Biol Soc. 2024;18(3):1-4.
68. Nirmala Bai L, Devi M. Pathogen-driven infectious disease recognition and classification: an in-depth analysis using machine learning and deep learning methods. Proc IEEE Int Conf Comput Sustain Glob Dev. 2024;1(1):1729-1734.
69. Zhang Y, Jiang H, Ye T, Juhas M. Deep learning for imaging and detection of microorganisms. Trends Microbiol. 2021;29(7):569-572.
70. Maaskant A, Lee D, Ngo H, Montijn RC, Bakker J, Langermans JAM, Levin E. AI for rapid identification of major butyrate-producing bacteria in rhesus macaques (Macaca mulatta). Anim Microbiome. 2025;7(1):39-54.

How to Cite: Urmila TS, Christy Eunaicy JI, Jayparatha C, Anandakumar JN, Govindaprabhu GB, Priya RM. Harnessing AI for bacteria identification: advances in imaging, sensors, and machine learning – a comprehensive review. Int Res J Multidiscip Scope. 2025; 6(4):950-977. doi: 10.47857/irjms.2025.v06i04.06174