INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Paper

Advances in Bacterial Classification: From Phenotypic Traits to Genomic Signatures

Ananya Das, Rojina Khatun, Sudeshna Sengupta, Malavika Bhattacharya*

Department of Biotechnology, Techno India University, West Bengal, EM-4, Sector-V, Salt Lake, Kolkata-700091, West Bengal, India

ARTICLE INFO

Published: 08 May. 2025

Keywords:

Bacterial classification, cell imaging, CNNs, DNNs

DOI:

10.5281/zenodo.15367325

ABSTRACT

Our knowledge of the biology of bacterial systems and classification methods has greatly increased as a result of recent developments in bacterial classification using cell imaging. In other instances, imaging has fueled fascinating developments in bacterial cell biology that have resulted in a finer knowledge of the mechanisms behind protein localization and cell growth. Geometric elements taken from digital microscopic pictures have been used to create automated systems for the identification and classification of bacterial cells. The current study's goal is to create an automated system for recognizing and categorizing bacterial cells in digital microscopic cell photographs. Bacilli, cocci, and spirilla are the three forms of bacterial cells that are distinguished by their geometric properties. The existing techniques depend on a human expert's subjective interpretation of profiles using a variety of hand-staining techniques. These techniques have demonstrated potential in the classification of spiral bacteria, including Vibrio, Spirillum, and spirochete, as well as bacilli and cocci. More recently, the classification of bacterial images has shown remarkable accuracy because of deep learning techniques, especially Convolutional Neural Networks (CNNs). Bacterial pictures were classified into 20 medically important groups with 99.9% accuracy using a ResNet-50 CNN model. Clinical experts have traditionally classified using traditional methods that do not rely on prediction tools. Manually classifying germs is a difficult and time-consuming process that takes a lot of human labor.The study employs Convolutional Neural Network (CNN) and Deep Neural Network (DNN) for image classification, a promising machine-learning technique.

***Corresponding Author:** Malavika Bhattacharya

Address: *Department of Biotechnology, Techno India University, West Bengal, EM-4, Sector-V, Salt Lake, Kolkata-700091, West Bengal, India*

Email ✉: malavikab@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

INTRODUCTION

Taxonomy is an organizing principle of biology and is ideally based on evolutionary relationships among organisms. The development of a robust bacterial taxonomy has been hindered by an inability to obtain most bacteria in pure culture and, to a lesser extent, by the historical use of phenotypes to guide classification. Bacterial classification is based on microbial techniques such as isolation, 6 culture, and staining to determine morphological and biochemical features. Over the years, in 7 addition to morphological characters, phenotypic and genotypic methods have gained much 8 importance in bacterial identification and classification schema. Recently, metabarcoding and metagenomics have become robust tools for studying bacterial and archaeal diversity by comparative genomic analysis. Increased reliability, decreased costs, and time for whole-genome sequencing and assembly have led to an unprecedented amount of information that can provide significant insights intodrugrepositioning and the development of robust molecular diagnostic platforms [1]. Molecular marker-based identification methods are widely used for the identification and classification of microbes from diverse environments. Over the past several decades, several molecular markers have been used for the identification and phylogenetic classification of specific microbial taxa [8–12]. The species selection criteria were i) genome assemblies must be in the ‘complete’ stage, ii) availability of at least five genome assemblies for a given species, and iii) genome assemblies (excluding plasmids) should have a minimum 95% average nucleotide identity (gANI) score at the intra-species level. Since the 1960s, phenotypic methods have been complemented by genetic methods including DNA-DNA hybridization, 16S rRNA sequence analysis, DNA G+C content, multilocus sequence typing (MLST), and multilocus sequence analysis (MLSA). DNA-DNA hybridization has been the gold standard procedure in bacterial taxonomy. Isolates showing >70 % DNA-DNA hybridization and differences in their melting temperature (ΔTm) of <5 degreesCelsius were considered to belong to the same species [2]. Bacterial classification has evolved from traditional phenotypic methods to more advanced automated techniques. Early approaches relied on subjective groupings and phenotypic similarities; modern methods incorporate genetic and molecular data for a polyphasic approach. Recent advancements include automated deep learning systems, such as the ResNet-50 CNN architecture, which achieved 99.2% accuracy in classifying bacterial images into 33 categories. Other machine learning algorithms, like support vector machines and random forests, have also been applied to bacterial classification. While DNA-DNA hybridization and ribosomal RNA analysis have improved taxonomic. Challenges remain in reconciling phylogenetic and practical classification systems. Despite concerns about lateral gene transfer, many well-defined genotypic clusters still correspond to species delineated by polyphasic approaches. Automated classification methods offer promising solutions for rapid and accurate bacterial identification in various fields [3].

Traditional phenotypic methods in bacterial classification

Traditional phenotypic methods for bacterial identification and typing, such as biotyping, antibiogram typing, and biochemical profiling, have been widely used but often yield erroneous or inconclusive results. These methods are labor-intensive, time-consuming, and less accurate compared to molecular techniques. While phenotypic methods can still be useful for epidemiological investigations of certain pathogens like Staphylococcus aureus, their results are heavily influenced by environmental conditions. Molecular methods, particularly 16S rRNA sequencing, have shown superior accuracy, speed, and ability to identify novel taxa compared to traditional phenotypic approaches. The advent of high-throughput phenotyping techniques has addressed some limitations of traditional methods, enabling rapid, non-destructive, and large-scale data acquisition. Overall, genotypic methods are considered more reliable, reproducible, and discriminatory than phenotypic techniques for bacterial typing and identification [4].

Traditional genotypic methods in bacterial classification

Bacterial classification has evolved from traditional phenotypic methods to a comprehensive polyphasic approach, integrating genotypic, chemotaxonomic, and phenotypic data. This approach includes 16S rRNA gene sequencing, DNA-DNA hybridization, molecular markers analysis, and biochemical tests. Genotype-based methods have become crucial, encompassing single-gene sequencing, molecular fingerprinting, and whole-genome analysis. Multilocus sequence typing (MLST) and analysis (MLSA) are powerful tools for species-level taxonomy. While some argue that lateral gene transfer complicates classification, many well-defined genotypic clusters still align with species delineated by polyphasic approaches. The ultimate goal is a theory-based classification system founded on phylogenetic and evolutionary concepts. However, the field remains dynamic, with classification schemes evolving alongside new techniques and approaches [5].

Subjective groupings of bacterial classification

Bacterial classification has evolved from subjective groupings to more objective systems based on phenotypic and genetic similarities. Early classifications relied on empirical principles, such as grouping bacteria by shared properties like denitrification or flower color. However, these groupings were often heterogeneous and lacked stability. The field progressed through various approaches, including numerical taxonomy, chemotaxonomic methods, and genetic analysis [6]. Currently, a polyphasic approach combining phenotypic, chemotaxonomic, and genetic data is widely used. Despite advancements, some argue that traditional taxonomic principles may not be suitable for bacteria, suggesting the need for different, potentially "heretical" approaches. The ultimate goal is to achieve a theory-based classification system rooted in phylogenetic and evolutionary concepts, although challenges remain due to factors like lateral genes. The classification of bacteria has evolved from subjective groupings to more objective systems based on phenotypic similarities and genetic relatedness. Classifications based on phenotypic characters lack stability, whereas those based on genetic relatedness tend to be stable. DNA-DNA hybridization has proven to be extremely useful in resolving taxonomic problems at the species level. Early taxonomic approaches were often empirical and practical, such as grouping bacteria by a single shared property. Biological classifications, the classifications of organisms, are tools of basic importance in research. Classifications not only provide some kind of order out of chaos but also facilitate identification, which makes the exchange of information possible and reduces unnecessary repetition of experiments. However, these classifications lacked stability and failed to reflect natural relationships [7]. According to definitions given by Simpson in 1961, taxonomy deals with the bases, principles, procedures, and rules of classification, and classification is the process of grouping objects according to their relationships. Bacteria can be classified based on a single property shared by all members of each group: denitrifiers, nitrogen fixers, hydrogen oxidizers, and plant pathogens. The development of numerical taxonomy improved phenotypic identification but provided limited phylogenetic information. Modern bacterial classification employs a polyphasic approach, integrating phenotypic, chemotaxonomic, and genotypic data with phylogenetic information. Despite advances, challenges remain in achieving a theory-based classification system. Some researchers argue that lateral gene transfer complicates the determination of evolutionary relationships, while others maintain that reliable classification is still possible using core genes and character genes as phylogenetic markers. The debate highlights the ongoing tension between practical and phylogenetic approaches to bacterial taxonomy [8].

Phenotypic similarities

Bacterial classification has evolved from subjective groupings to more objective systems based on phenotypic similarities and genetic relatedness. While DNA-DNA hybridization and 16S rRNA gene sequencing have proven useful for species-level identification, phenotypic characteristics remain crucial in resolving conflicts and complementing genetic data. The current classification of Bacteria and Archaea employs a polyphasic approach, integrating phenotypic, chemotaxonomic, and genotypic data with phylogenetic information [9]. Despite debates on the impact of lateral gene transfer on classification, many well-defined genotypic clusters align with species delineated by polyphasic approaches. Comparative sequence analysis of core genes may be useful for characterizing higher taxa, while various character genes could serve as phylogenetic markers for lower taxa. Ultimately, scientific classification aims to describe overall similarities between organisms, with similarity best measured by shared features and taxa division based on correlated features. Broad relationships among bacteria have been identified by comparing ribosomal RNA cistrons; however, many groups based on ribosomal RNA analysis are not easily definable in terms of phenotypic similarities. Classification of bacteria evolved from limited subjective groupings to general, more objective arrangements based on overall phenotypic similarities. However, classifications based on phenotypic characters lack stability, whereas those based on genetic relatedness tend to be stable. Phenotypic characteristics, while complementing 16S rRNA gene sequencing in identifying bacteria, become decisive in solving conflicts of equal % similarity of a given DNA sequence to more than one classified [10]. Predictions of microbial phenotypes from metagenomic data depend heavily on our knowledge of expressed genes. Thus renaissance of microbial phenotypic characterization is likely to emerge at par with genotypic signatures. Many schemes of bacterial taxonomy are not classifications but catalogs. The similarity is best measured by the number of features in common between two strains, while division into taxa is based on correlated Other criteria for these two basic ideas are unsatisfactory and confusing, and there seems to be no logical reason why any one feature should be given greater weight in classification than any other. Hierarchical systems are a practical necessity, and simple mathematical methods are useful in bacterial classification. It is not necessary to know the evolutionary history of organisms to classify them. An attempt has been made to realize the reader's importance of the interplay of genotypic and phenotypic characteristics of bacteria for the development of comprehensive and more stable classification schemes. It is expected that future valid classification schemes will be based on the phenetic relationships of microorganisms [11].

Traditional morphological methods are insufficient for accurate classification.

Recent research highlights the limitations of traditional morphological methods for accurate classification in various biological and medical domains. Demonstrated that conventional deep CNNs and object recognition techniques are insufficient for distinguishing subtle visual differences in cellular colonies.Found that classical neural networks often result in suboptimal class separation and poor endmember distinction in hyperspectral signature classification. To address these challenges, advanced techniques have been proposed. Introduced a high-dimensional shape transformation method combined with a nonlinear SVM for brain image classification, achieving high accuracy in gender and age classification. Laosai&Chamnongthai]. Developed a coarse-to-fine approach for acute leukemia subtype classification, integrating morphological features with CD markers. These studies emphasize the need for more sophisticated methods, such as Triplet-net CNN learning [12].

Genetic and molecular data for a polyphasic approach in bacterial classification

The polyphasic approach to bacterial classification combines phenotypic, genotypic, and chemotaxonomic methods to provide a comprehensive taxonomic framework. This approach has become increasingly important as traditional morphological and biochemical methods alone are insufficient for accurate classification. Recent advances in molecular techniques, including 16S rRNA sequencing and DNA-DNA hybridization, have significantly improved our ability to delineate bacterial taxa. However, the increasing availability of whole-genome sequences offers new opportunities for taxonomic classification. Propose a "taxonomy-genomics" strategy that incorporates genomic data into the polyphasic approach, providing more reliable and reproducible results. Despite these advancements, challenges remain, such as the classification of unculturable bacteria and the potential impact of lateral gene transfer on phylogenetic relationships [13]. Nevertheless, the polyphasic approach remains the most comprehensive method for bacterial classification, allowing for a more accurate representation of microbial diversity. Polyphasic taxonomy has emerged as a comprehensive approach to bacterial classification, integrating phenotypic, genotypic, and chemotaxonomic data. This method combines traditional microbiological techniques with modern molecular methods, including 16S rRNA gene sequencing, DNA-DNA hybridization, and biochemical assays]. The polyphasic approach has become increasingly popular due to advances in molecular biology and bioinformatics, allowing for more accurate taxonomic placement. Recent developments have incorporated genomic data into the polyphasic strategy, a method termed "taxon-genomics". This approach utilizes whole-genome sequences to provide reliable and reproducible data for taxonomic classification. Despite challenges such as lateral gene transfer, the polyphasic approach continues to be effective in maintaining typical gene- and phenotypic characteristics of taxa for reliable classification and identification of Bacteria [14].

Automated DL systems for bacterial classification

Recent studies have demonstrated the effectiveness of automated deep learning systems, particularly using the ResNet-50 CNN architecture, for bacterial image classification. These systems achieve high accuracy rates, ranging from 99.2% to 99.9%, surpassing traditional manual methods. The ResNet-50 model, often implemented with transfer learning techniques, has shown success in classifying bacteria into multiple categories, with some studies examining up to 33 different types. This approach offers advantages such as reduced processing time, lower costs, and increased accuracy compared to conventional visual observations. Additionally, researchers have explored the potential of augmented reality to enhance the classification process and expand datasets for improved performance. These automated systems present a promising solution for efficient and accurate bacterial classification in clinical microbiology [15].

ML algorithms for bacterial classification

Machine learning algorithms have shown promising results in bacterial classification and antimicrobial susceptibility testing using MALDI-TOF mass spectra. Various supervised learning techniques, including Support Vector Machines, Genetic Algorithms, and Artificial Neural Networks, have been employed for microbial identification. Convolutional Neural Networks, particularly pre-trained models using transfer learning, have demonstrated high accuracy in classifying bacterial species from images, with DenseNet-121 achieving 99.08% accuracy on a dataset of 33 bacterial species. These machine-learning approaches offer advantages over traditional methods like Gram staining and biochemical testing. However, challenges remain, such as the need for external validation of algorithms and addressing shortcomings to incorporate them into clinical routines. Overall, machine learning techniques show great potential for enhancing healthcare methodologies and infectious disease prevention practices [16].

Evaluation of automatic bacterial classification methods and approaches

Automatic bacterial classification methods have gained prominence in microbiology, offering rapid and accurate identification of bacterial species. These approaches utilize machine learning algorithms, including support vector machines, deep learning, and random forests, to analyze data from various sources. Geometric features extracted from digital microscopic images can be used to classify bacteria into different types, such as bacilli, cocci, and spirilla. However, the accuracy of these methods varies widely, and no single program consistently outperforms others across all scenarios [17]. Evaluation of classification methods should include clade exclusion to better assess performance when identical sequences are not present in reference databases. Despite challenges, automated bacterial classification has the potential to revolutionize microbiology, offering high-throughput solutions for taxonomy, epidemiology, and microbial ecology. Traditional methods of bacterial classification have been labour-intensive and time-consuming, but recent advancements in automation and data-driven techniques have paved the way for automatic classification, significantly accelerating the process and enhancing its accuracy. Key components of automatic bacterial classification include data acquisition, feature extraction, model selection, training, and validation. For the automatic classification of this data, machine learning and artificial intelligence algorithms, including support vector machines, deep learning, and random forest models, were used. The objective of the present study is to develop an automatic tool to identify and classify bacterial cells in digital microscopic cell images[18]. We propose a method for bacterial classification by segmenting digital bacterial cell images and extracting geometric features for cell classification. Automatic bacterial classification methods have gained significant attention due to their potential to revolutionize microbiology by offering rapid and accurate identification of bacterial species. These methods typically involve image segmentation, feature extraction, and machine learning algorithms to classify bacteria based on their morphology. Convolutional Neural Networks (CNNs) have shown promising results in bacterial strain classification, achieving high accuracy and precision in distinguishing between different strains. Various geometric features and discriminative attributes are used to differentiate bacterial types, such as bacilli, cocci, and spiral shapes. These automated approaches offer advantages over traditional manual methods, including faster processing times and reduced reliance on experts[19].The integration of automatic classification techniques has the potential to enhance bacterial taxonomy, epidemiology, and microbial ecology.As the traditional methods of classification have numerous drawbacks This analysis explores into the realm of automatic classification of bacteria, exploring the innovative technologies and approaches that enable the rapid and precise categorization of bacterial species this analysis explores into the realm of automatic classification of bacteria, exploring the innovative technologies and approaches that enable the rapid and precise categorization of bacterial[19]. The integration of automatic classification of bacteria has the potential to revolutionize microbiology, offering rapid and high-throughput solutions for taxonomy, epidemiology, and microbial ecology [20]. It enables the real-time identification of bacterial species, facilitating quicker response to infectious diseases and more effective management of microbial communities in various environments. A method for bacterial classification by segmenting digital bacterial cell images and extracting geometric features for cell classification. The experimental results are compared with the manual results obtained by the microbiology expert and demonstrate the efficacy of the proposed method [21].

Automatic classification of bacterial cells in digital microscopic images

Automatic classification of bacterial cells in digital microscopic images has been a focus of recent research, aiming to develop reliable computer-assisted methods for analyzing microbial communities at single-cell resolution. Several studies have proposed approaches using geometric features to identify and classify different types of bacterial cells, including bacilli, cocci, and spiral bacteria. These methods involve segmenting digital bacterial cell images and extracting geometric features for classification using various algorithms such as k-NN, neural networks, and fuzzy classifiers [22]. Additionally, researchers have explored the classification of bacterial growth phases in bacilli cells. More recent work has employed advanced machine learning techniques, such as the Bag-of-Words model and Support Vector Machine (SVM) classifiers, achieving high accuracy in classifying different bacterial species from digital microscope images. These automated approaches aim to improve upon traditional manual methods, which rely on subjective interpretation by human experts. The objective of the study is to develop an automatic tool to identify and classify bacterial cells in digital microscopic cell images [23]. Geometric features are used to identify the different types of spiral bacterial cells, namely, vibrio, spirillium, and spirochete. The biologist's interpretation suffers from insufficient information and thus may lead to limited accuracy in the bacteria classification process. To handle this drawback, machine learning tools and image analysis approaches have tackled the identification of different bacteria species for improving clinical microbiology. The current methods rely on the subjective reading of profiles by a human expert based on the various manual staining methods [24].

Evaluation of automatic bacterial classification methods and approaches.

Automatic classification methods using machine learning and AI can enable rapid and accurate identification of bacterial species. Accurate and efficient classification of bacteria is essential for understanding their ecological roles, diagnosing infectious diseases, and unlocking the potential of beneficial microorganisms in biotechnology. Advancements in automation and data-driven techniques have eased the way of classification and, most importantly, sped up the process with higher accuracy [25]. The integration of automatic classification of bacteria has the potential to revolutionize microbiology, offering rapid and high-throughput solutions for taxonomy, epidemiology, and microbial ecology. It enables the real-time identification of bacterial species, facilitating quicker response to infectious diseases and more effective management of microbial communities in various settings. Feature Extraction is the process of identifying and selecting important characteristics or patterns from raw data to create a more simplified and meaningful representation for further analysis or processing [26]. Because of the pathogenicity of various bacterial species and the necessity to characterize the ecologically and economically beneficial species, bacterial strains are classified at the species level. ML techniques are widely employed by researchers for studying different bacterial species. In the year 1998, Veropoulos et al. proposed an artificial neural network (ANN) based technique for the automatic image recognition of tuberculosis bacteria in Ziehl–Neelsen (ZN) stained sputum smear images. The dataset used for the study included 267 bacillus images and 88 non-bacillus images. The methodology involved edge-based segmentation using a Canny edge detector, followed by shape features extraction using a discrete Fourier transform. Automated recognition and classification of bacteria species from microscopic images have significant importance in clinical microbiology. The proposed method achieved an average classification accuracy of 99.2%. The experimental results demonstrate that the proposed technique surpasses state-of-the-art methods in the literature and can be used for any type of bacteria classification task. Key components of automatic bacterial classification include data acquisition, feature extraction, model selection, training, and validation [27]. For the automatic classification of this data, machine learning and artificial intelligence algorithms, including support vector machines, deep learning, and random forest models, were used. Automatic bacterial classification methods have gained significant attention due to their potential to revolutionize microbiology. These approaches employ various machine learning algorithms, including support vector machines, deep learning, and random forests, to rapidly categorize bacterial species. Evaluation of these methods has revealed wide variability in sensitivity, precision, and computational demands. Notably, clade exclusion techniques have been used to assess method performance under progressively challenging scenarios [28]. Machine learning approaches have shown tremendous performance in automatic bacterial detection, assisting microbiologists in solving complex problems. Geometric features extracted from digital microscopic images have been utilized to identify and classify different bacterial cell types, such as bacilli, cocci, and spiral. While these automated methods offer promising solutions, researchers must carefully consider each method's strengths and limitations to avoid misleading results. Accurate and efficient classification of bacteria is essential for understanding their ecological roles, diagnosing infectious diseases, and unlocking the potential of beneficial microorganisms in biotechnology [29. While traditional methods of bacterial classification have been labor-intensive and time-consuming, recent advancements in automation and data-driven techniques have paved the way for automatic classification, significantly accelerating the process and enhancing its accuracy. This analysis explores the realm of automatic classification of bacteria, exploring the innovative technologies and approaches that enable the rapid and precise categorization of bacterial species. Key components of automatic bacterial classification include data acquisition, feature extraction, model selection, training, and validation. For the automatic classification of this data, machine learning and artificial intelligence algorithms, including support vector machines, deep learning, and random forest models, were used [30]. The field of metagenomics (the study of genetic material recovered directly from an environment) has grown rapidly, with many bioinformatics analysis methods being developed. For taxonomic classification of sequence reads, such evaluation should include the use of clade exclusion, which better evaluates a method's accuracy when identical sequences are not present in any reference database, as is common in metagenomic analysis. An overview of the features of 38 bioinformatics methods is provided, evaluating accuracy with a focus on 11 programs that have reference databases. Taxonomic classification of sequence reads was evaluated using both in silico and in vitro mock bacterial communities [31]. Clade exclusion was used at taxonomic levels from species to identify how well methods perform in progressively more difficult scenarios.A wide range of variability was found in the sensitivity, precision, overall accuracy, and computational demand for the programs evaluated. In experiments where distilled water was spiked with only 11 bacterial species, frequently dozens to hundreds of species were falsely predicted by the most popular programs. The different features of each method (forces predictions or not, etc.) are summarized, and additional analysis considerations are discussed. The accuracy of shotgun metagenomics classification methods varies widely. The accuracy of shotgun metagenomics classification methods varies widely. No one program outperformed others in all evaluation scenarios; rather, the results illustrate the strengths of different methods for different purposes. Researchers must appreciate method differences, choosing the program best suited for their particular analysis to avoid very misleading results[32].The use of standardized datasets for method comparisons is encouraged, as is the use of mock microbial community controls suitable for a particular metagenomic analysis.Forthe automatic classification of this data, machine learning and artificial intelligence algorithms, including support vector machines, deep learning, and random forest models, were used. The integration of automatic classification of bacteria has the potential to revolutionize microbiology, offering rapid and high-throughput solutions for taxonomy, epidemiology, and microbial ecology [33]. Bacteria are important in a variety of practical domains, including industry, agriculture, and medicine. etc. Very few species of bacteria are favorable to humans. Whereas, the majority of them are extremely dangerous and cause a variety of life-threatening illnesses to different living beings. Traditionally, this class of microbes is detected and classified using different approaches, like Gram staining, biochemical testing, motility testing, etc. However, with the availability of large amounts of data and technical advances in the field of medical and computer science, machine learning methods have been widely used and have shown tremendous performance in the automatic detection of bacteria [34].

Automatic identification and classification

Recent studies have explored automated methods for identifying and classifying bacterial cell growth phases using digital microscopy and spectroscopy. Hiremath and Bannigidad (2010) developed an automatic tool to classify bacilli bacterial cells into normal, grown-up, and about-to-divide phases using geometric features and various classifiers. They extended this work to classify different bacterial cell types (bacilli, cocci, and spirilla) based on geometric features. The process of recognizing bacteria based solely on their shape would be a difficult one because many bacteria share very similar shapes [35]. The second most differentiating feature is the shape and size of the colonies formed by the bacteria. A different approach used Fourier transform infrared spectroscopy to detect biochemical changes in bacterial cells during different growth phases. They observed major variations in cellular structure across growth phases for E. coli and L. innocua, demonstrating the potential of spectroscopic techniques for bacterial growth. The study examined the potential of Fourier transform infrared (FT-IR) absorbance spectroscopy to detect biochemical changes in bacterial cells that occur during bacterial growth phases in batch culture [36]. A method is proposed for bacterial cell classification based on their different growth phases by segmenting digital bacilli bacterial cell images and extracting geometric features for cell growth phase identification and classification using 3 classifiers: k-NN classifier, Neural Network classifiers, and Fuzzy classifiers. The focus of this study was to examine Fourier transform infrared (FT-IR) spectral features of bacterial cells at different growth stages and to determine if variations in composition and distribution of the biochemical components of cells during growth phases can be distinguished using this spectroscopic technique. Major variations in the biochemical structure of bacterial cells during growth were observed. Over the range of 1800–900 cm^{-1}, loadings 1 (principal component [PC] 1) and 2 (PC2) accounted for 88% of the total variability (76% and 12%, respectively) for E. coli cells, and 80% (72% and 8%, respectively) for L. innocua cells. Recent studies have explored deep-learning approaches for automated bacterial classification from microscopic images [37]. Convolutional Neural Networks (CNNs) have shown promising results, with pre-trained architectures like ResNet-50 achieving 99.2% accuracy across 33 bacteria. Transfer learning techniques have been employed to enhance classification performance and accelerate training. Combining deep learning feature extraction with Support Vector Machine (SVM) classification has yielded impressive results, with 99.61% accuracy reported for bacterial colony classification. These automated approaches aim to reduce analysis time and improve diagnostic accuracy compared to manual methods. While most studies focused on multiple bacterial genera, even limited datasets of two genera have shown potential for successful classification using deep learning techniques. An automated process for bacteria recognition becomes attractive to reduce the analysis time and increase the accuracy of the diagnostic process [38]. This research studiesthe possibility of using image classification and deep learning methodsto classify genera of bacteria. Automated recognition and classification of bacteria species from microscopic images have significant importance in clinical microbiology. Bacteria classification is usually carried out manually by biologists using different shapes and morphologic characteristics of bacteria species]. The manual taxonomy of bacteria types from microscopy images is time-consuming and a challenging task for even experienced biologists. In this study, an automated deep learning-based classification approach has been proposed to classify bacterial images into different categories. Rapid diagnosis is a significant issue in improvingthe quality of bacterial detection. Current computer-aided diagnosis (CAD) based image processing and machine learning proposed for bacteria in clinical microbiology [39]. CAD utilized several machine learning (ML) approaches to classify the various bacterial species that used some classifiers such as support vector machine (SVM), backpropagation neural network (BPNN), real AdaBoost, and modest AdaBoost. The implementation results have confirmed that images from a microscope can recognize the genus of the bacterium. The experimental results compare the deep learning methodology for accuracy in bacteria recognition, standard resolution imagesin the use case study. The research studies the possibility of using image classification and deep learning methodsto classify genera of bacteria. We propose the implementation method of the bacteria recognition system using Python programming and the Keras API with the TensorFlow Machine Learning framework [40]. The ResNet-50 pre-trained CNN architecture has been used to classify digital bacteria images into 33 categories. The transfer learning technique was employed to accelerate the training process of the network and improve its classification performance of the network. The proposed method achieved an average classification accuracy of 99.2%. The experimental results demonstrate that the proposed technique surpasses state-of-the-art methods in the literature and can be used for any type of bacteria classification task. Transfer learning techniques have been employed to enhance model performance, with GoogLeNet reaching 98.67% accuracy. A comprehensive framework combining Convolutional Deep Belief Networks for segmentation and CNNs for classification has been proposed, outperforming classical methods [41]. Bacterial image segmentation and classification an important problem because bacterial appearance can vary dramatically based on environmental conditions. Further, newly isolated species may exhibit phenotypes previously unseen. The DIBaS dataset, containing 660 images of 33 bacterial genera and species, has been introduced to facilitate comparative studies. The ability to identify and categorize bacteria is crucial in modern medicine for disease diagnosis, infection treatment, and epidemic investigation. However, manual identification and categorization of bacteria require a lot of time and effort from humans [42].

CONCLUSION

In conclusion, the dynamic field of bacterial classification stands as a cornerstone of microbiological understanding, constantly evolving in response to technological advancements and our deepening appreciation of microbial diversity. From the foundational phenotypic methods to the revolutionary impact of molecular techniques, particularly whole-genome sequencing, and phylogenomic analyses, our ability to delineate and categorize bacteria has undergone a profound transformation. These advancements have not only refined our taxonomic frameworks, revealing intricate evolutionary relationships and previously cryptic species, but have also provided crucial insights into bacterial ecology, pathogenicity, and biotechnological potential. While current polyphasic approaches, integrating phenotypic, genotypic, and chemotaxonomic data, offer a robust and increasingly accurate system, challenges remain. The sheer scale of microbial diversity, the complexities of horizontal gene transfer, and the ongoing need for standardized and accessible methodologies necessitate continued research and collaborative efforts within the scientific community. Future directions will likely focus on leveraging artificial intelligence and machine learning to analyze vast genomic datasets, developing more sophisticated methods for culturing and characterizing unculturable bacteria and establishing more universally accepted guidelines for taxonomic revisions. Ultimately, a refined and comprehensive bacterial classification system is not merely an academic pursuit; it is fundamental to addressing global challenges in healthcare, agriculture, and environmental sustainability, paving the way for a deeper understanding of the microbial world that underpins life on Earth.

ACKNOWLEDGMENT

Firstly, I extend my heartfelt gratitude to Dr. Malavika Bhattacharya, Sudeshna Sengupta, and Rojina Khatun for their invaluable guidance, encouragement, and support throughout the project. Their expertise and constructive feedback played a vital role in shaping the direction and outcome of my work.

REFERENCES

1. Parks DH, Chuvashia M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018.
2. Paul B, Dixit G, Murali TS, Satyamoorthy K. Genome-based taxonomic classification. Genome. 2019;62(2):45–52.
3. Paul B, Kavia Raj K, Murali TS, Satyamoorthy K. Species-specific genomic sequences for classification of bacteria. Comput Biol Med. 2020; 123:103874.
4. Margos G, Fingerle V, Cutler S, Gofton A, Stevenson B, Estrada-Peña A. Controversies in bacterial taxonomy: The example of the genus Lorelai. Ticks Tick Borne Dis. 2020;11(2):101335.
5. Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol. 2009;32(8):533–42.
6. Talo M. An Automated Deep Learning Approach for Bacterial Image Classification. arXiv. 2019; abs/1912.08765.
7. Keren F, Keziah F, Fredrick Gnanaraj F, Vanitha L. Evaluation of Automatic Bacterial Classification Methods and Approaches. In: 2023 9th Int Conf on Smart Structures and Systems (ICSSS). 2023. p.1–6.
8. Kasela M, Malm A. Overview of phenotypic methods used for differentiation of Staphylococcus aureus. Curr Issues Pharm Med Sci. 2018; 31:117–21.
9. Xiao Q, Bai X, Zhang C, He Y. Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review. J Adv Res. 2021; 35:215–30.
10. Ramadan AA. Bacterial typing methods from past to present: A comprehensive overview. Gene Rep. 2022.
11. Vandamme PA. Taxonomy and classification of bacteria. 2015.
12. Prakash O, Verma M, Sharma P, Kumar MJ, Kumari K, Singh AK, et al. Polyphasic approach of bacterial classification — an overview of recent advances. Indian J Microbiol. 2007; 47:98–108.
13. Felis GE, Torriani S, Vlieg JE, Oren A. Taxonomic Characterization of Prokaryotic Microorganisms. 2010.
14. Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol. 2009;32(8):533–42.
15. Felis GE, Torriani S, Vlieg JE, Oren A. Taxonomic Characterization of Prokaryotic Microorganisms. 2010.
16. Palleroni NJ. The Taxonomy of Bacteria. BioScience. 1983; 33:370–7.

17. Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol. 2009;32(8):533–42.
18. Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol. 2009;32(8):533–42.
19. Tabssum F, Ahmad Q, Qazi JI. DNA sequenced based bacterial taxonomy should entail decisive phenotypic remarks: Towards a balanced approach. J Basic Microbiol. 2018; 58:918–27.
20. Witmer A, Theagarajan R, Bhanu B. Triplet-Net Classification of Contiguous Stem Cell Microscopy Images. IEEE/ACM Trans Comput Biol Bioinform. 2023; 20:2314–27.
21. Schmalz MS, Ritter GX. Hyperspectral endmember extraction and signature classification with morphological networks. 2006.
22. Lao Z, Shen D, Xue Z, Karaçali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. Neuroimage. 2004; 21:46–57.
23. Laosai J, Chamnongthai K. Classification of acute leukemia using medical-knowledge-based morphology and CD marker. Biomed Signal Process Control. 2018; 44:127–37.
24. Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol. 2009;32(8):533–42.
25. Das S, Dash HR, Mangwani N, Chakraborty J, Kumari S. Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. J Microbiol Methods. 2014; 103:80–100.
26. Prakash O, Verma M, Sharma P, Kumar MJ, Kumari K, Singh AK, et al. Polyphasic approach of bacterial classification—An overview of recent advances. Indian J Microbiol. 2007; 47:98–108.
27. Ramasamy D, Mishra AK, Lagier J, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J SystEvolMicrobiol. 2014;64(Pt 2):384–91.
28. Prakash O, Verma M, Sharma P, Kumar MJ, Kumari K, Singh AK, et al. Polyphasic approach of bacterial classification — An overview of recent advances. Indian J Microbiol. 2007; 47:98–108.
29. Talo M. An Automated Deep Learning Approach for Bacterial Image Classification. arXiv. 2019; abs/1912.08765.
30. Wonohadidjojo DM. Classification of Bacterial Images using Transfer Learning, Optimized Training and Resnet-50. Eduvest J Univ Stud. 2022.
31. Nagro SA, Alansari MM, Sendy BK, Abutarboush MH. Microscopic Bacterial Classification Using Deep Learning Based on Augmented Reality Architecture. In: 2023 3rd International Conference on Computing and Information Technology (ICCIT). 2023. p. 370–6.
32. Kotwal S, Rani P, Arif T, Manhas JS, Sharma S. Automated Bacterial Classifications Using Machine Learning Based Computational Techniques: Architectures, Challenges and Open Research Issues. Arch Comput Methods Eng. 2021; 29:2469–90.
33. Honeine P, Sessa PG, Zürich E, Shaaban S, Sahmoud, Wu Y, et al. Machine learning algorithms in microbial classification: a comparative analysis. Front ArtifIntell. 2023;6.
34. Weis C, Cr J, K B. Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. Clin Microbiol Infect. 2020.

35. Keren F, Keziah F, Fredrick Gnanaraj F, Vanitha L. Evaluation of Automatic Bacterial Classification Methods and Approaches. In: 2023 9th International Conference on Smart Structures and Systems (ICSSS). 2023. p. 1–6.
36. Peabody MA, Van Rossum T, Lo R, Brinkman FS. Evaluation of shotgun metagenomics sequence classification methods using in silico and in vitro simulated communities. BMC Bioinformatics. 2015;16.
37. Hiremath PS, Bannigidad P. Automatic classification of bacterial cells in digital microscopic images. In: International Conference on Digital Image Processing. 2010.
38. Trivedi S, Patel N, Faruqui N. Bacterial Strain Classification using Convolutional Neural Network for Automatic Bacterial Disease Diagnosis. In: 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence). 2023. p. 325–32.
39. Mohamed BA, Afify HM. Automated classification of Bacterial Images extracted from Digital Microscope via Bag of Words Model. In: 2018 9th Cairo International Biomedical Engineering Conference (CIBEC). 2018. p. 86–9.
40. Hiremath PS, Bannigidad P. Digital Microscopic Image Analysis of Spiral Bacterial Cell Groups. 2011.
41. Rani P, Kotwal S, Manhas JS, 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. 2021; 29:1801–37.
42. Talo M. An Automated Deep Learning Approach for Bacterial Image Classification. arXiv. 2019; abs/1912.08765.

HOW TO CITE: Ananya Das, Rojina Khatun, Sudeshna Sengupta, Malavika Bhattacharya*, Advances in Bacterial Classification: From Phenotypic Traits to Genomic Signatures, Int. J. of Pharm. Sci., 2025, Vol 3, Issue 5, 1293-1306. https://doi.org/10.5281/zenodo.15367325