

Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements

Syed Khasim¹, Hritwik Ghosh¹, Irfan Sadiq Rahat^{1,*}, Kareemulla Shaik¹, Manava Yesubabu²

¹School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh

²Department of Computer Science Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad, India

Abstract

Microorganisms are pervasive and have a significant impact in various fields such as healthcare, environmental monitoring, and biotechnology. Accurate classification and identification of microorganisms are crucial for professionals in diverse areas, including clinical microbiology, agriculture, and food production. Traditional methods for analyzing microorganisms, like culture techniques and manual microscopy, can be labor-intensive, expensive, and occasionally inadequate due to morphological similarities between different species. As a result, there is an increasing need for intelligent image recognition systems to automate microorganism classification procedures with minimal human involvement. In this paper, we present an in-depth analysis of ML and DL perspectives used for the precise recognition and classification of microorganism images, utilizing a dataset comprising eight distinct microorganism types: Spherical bacteria, Amoeba, Hydra, Paramecium, Rod bacteria, Spiral bacteria, Euglena and Yeast. We employed several ml algorithms including SVM, Random Forest, and KNN, as well as the deep learning algorithm CNN. Among these methods, the highest accuracy was achieved using the CNN approach. We delve into current techniques, challenges, and advancements, highlighting opportunities for further progress.

Keywords: Microorganisms, biotechnology, classification, clinical microbiology, food production, Amoeba, Euglena, Hydra, Paramecium, Rod bacteria, Spherical bacteria, Spiral bacteria, Yeast, SVM, Random Forest, KNN, CNN

Received on 23 September 2023, accepted on 20 November 2023, published on 27 November 2023

Copyright © 2023 S. Khasim *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetiot.4484

*Corresponding authors Email: me.rahat2020@gmail.com

1 Introduction

Microorganisms, the ubiquitous and diverse life forms, play an integral role in numerous fields, including healthcare, environmental monitoring, and biotechnology [11]. Their accurate identification and classification are essential for a variety of applications, including clinical microbiology, agriculture, and food production [12]. Microorganisms, for example, can be employed in bioremediation, biofertilizers, and biofuel generation. [13]. However, they also pose risks in the form of pathogenic organisms that cause infectious diseases [14].

Traditional methods for studying microorganisms, like culture techniques and manual microscopy, can be time-

consuming, costly, and occasionally inadequate due to morphological similarities between different species [15]. Consequently, the development of intelligent image recognition tools that can automate microorganism classification processes with minimal human intervention has become increasingly pertinent.

ML and DL advancements in recent years have demonstrated impressive achievements in a variety of application areas, including image recognition, object segmentation, pattern recognition, and autonomous vehicles [16]. Capitalizing on these accomplishments, researchers have started to explore the application of ML and DL methodologies for microorganism image recognition, aiming at species-level identification and classification. These techniques have been employed in image preprocessing,

feature extraction, and classification, significantly enhancing the efficiency and accuracy of microorganism analysis [17].

This research paper, titled "Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements," focuses on a specific dataset containing images of eight different microorganisms: Amoeba, Euglena, Hydra, Paramecium, Rod bacteria, Spherical bacteria, Spiral bacteria, and Yeast are some of the bacteria found in the environment. By employing ML and DL algorithms, our goal is to develop a robust and accurate classification system for these microorganisms. In the following sections, we provide an exhaustive review of ML and DL approaches employed in deciphering microorganisms through intelligent image recognition. We delve into the methodologies, challenges, and advancements in the field, focusing on the application of ML and DL techniques for microorganism image analysis, including image categorization, feature extraction, and segmentation are a few examples. Additionally, we discuss the challenges faced by researchers in this domain, covering image preprocessing, data augmentation, feature extraction, and model selection. Lastly, we highlight recent advancements in the field, emphasizing the implementation Transfer learning, generative adversarial networks (GANs) and unsupervised learning techniques are all examples of ml.

By presenting an in-depth analysis of the existing methodologies and their limitations, we aspire to inspire further investigation and collaboration, ultimately contributing to the development of effective, reliable, and intelligent systems for microorganism image. We present a paper focused on the topic of deciphering microorganisms through intelligent image recognition, utilizing machine learning and deep learning approaches. The content is targeted towards computer scientists and researchers in related fields, with examples of various microorganisms provided in [Fig. 1] to provide context and aid in understanding.

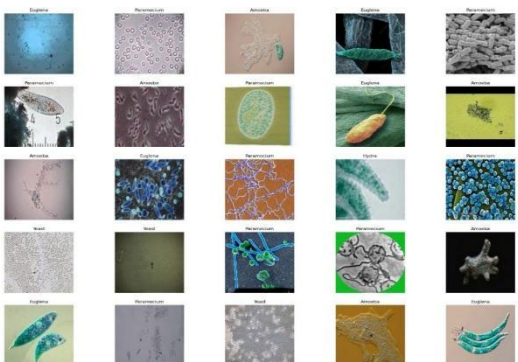


Fig 1. The examples of images of the investigated classes of microorganisms

1.1 The Advantages of Microorganisms

Microorganisms are a diverse group of living organisms that play an integral role in various fields such as healthcare, environmental monitoring, and biotechnology [1]. The accurate identification and classification of microorganisms are critical for different applications, including clinical microbiology, agriculture, and food production [2]. The advantages of microorganisms are many, including their ability to produce useful products like antibiotics, enzymes, and biofuels through biotechnology [3]. Microorganisms can also be employed in environmental monitoring, waste management, and food production [4]. In agriculture, microorganisms play a vital role in enhancing soil quality, promoting plant growth, and protecting crops from pests and diseases [5]. Therefore, it is essential to accurately identify and classify microorganisms for their diverse applications. Microorganisms have numerous benefits and play a crucial role in several fields, including biotechnology, environmental monitoring, and agriculture. Their accurate identification and classification are essential for various applications, including the production of useful products and the prevention of infectious diseases.

2 Literature Review

An extensive body of literature has been dedicated to the investigation of intelligent image recognition techniques for understanding microorganisms. Gray and colleagues (2002) [1] examined multiple image analysis approaches for estimating algal cell counts, comparing various segmentation techniques based on thresholding, edge detection, and template matching. Qiu and co-authors (2004) [2] chronicled the evolution of bacteria counting and cell size measurement methods, encompassing both traditional methodologies and automated flow analysis technologies. Gracias (2004) [3] addressed the application of fluorogenic or chromogenic strategies to differentiate between bacterial species and the use of impedance technology for enumeration.

The distinction between microbe counting and biovolume measurement made by Daims and Wagner in 2007 [4] placed emphasis on the identification or lack thereof of specific entities (cells or cell clusters) within biomass. The item counting techniques investigated by Barbedo (2012a) [5] included morphological operations, filtering, contrast augmentation, transformations, edge detection, and image segmentation. The use of CMEIAS for microbe counting and biovolume measurement through image processing was described by Dazzo and Niccum (2015) [6], who also included hierarchical tree classifiers and k-Nearest Neighbour classifiers for classification purposes.

Li and collaborators (2019a) [7] conducted a comprehensive review of the development of computer-based microorganism image analysis, presenting various classification methods for different microorganisms.

Puchkov (2019) [8] delineated the primary quantitative analysis approaches for single bacterial and yeast cells at the cellular and subcellular levels.

Evaluating the existing literature reveals that a variety of ML and DL techniques have been employed to address microorganism image recognition and classification. Nevertheless, there is still ample room for improvement concerning efficiency, accuracy, and adaptability to a variety of imaging conditions and sources.

Shi et al. (2018) suggested a binary matrix completion-based prediction method (BMCMDA) in their paper [9]. The incomplete microbiome-disease association (MDA) matrix is believed to be the sum of a possible parameterization matrix and a noise matrix, with the MDA matrix items' independent subscripts following a binomial model. BMCMDA exceeds KATZHMMA in terms of AUC and can be supplemented with additional independent microbial/disease similarities or traits to improve MDA prediction. This method can also be used to forecast a variety of other factors.

Fan et al. (2019) [10] established a new technique for assessing microbial-disease connections by merging data from the MDPH HMDA with path-based HeteSim scores. First, heterogeneous networks were constructed, followed by microbe-disease pair weighting.

3 Description of the Dataset

The Microorganism Classification Dataset represents a rich and valuable resource for researchers and practitioners working on the cutting edge of microbiology, machine learning, and deep learning. This meticulously curated dataset encompasses a total of 795 microscopic images spread across eight distinct categories. These categories include Amoeba (75 images), Euglena (170 images), Hydra (75 images), Paramecium (155 images), Rod bacteria (85 images), Spherical bacteria (85 images), Spiral bacteria (75 images), and Yeast (75 images). Each folder in the dataset contains high-quality images captured under controlled conditions, which ensures that the dataset effectively represents the morphological characteristics of each microorganism type. The variation in the number of images per category reflects the diversity and complexity of the microorganisms, making the dataset an ideal benchmark for developing and testing intelligent image recognition models. By providing such a comprehensive collection of microscopic images, the Microorganism Classification Dataset fosters a deeper understanding of microorganism classification and encourages the development of advanced models that can accurately and efficiently differentiate between various microorganism types. This dataset ultimately contributes to the broader goal of advancing microbiological research and its applications in fields such as medicine, agriculture, and environmental science.

3.1 Preprocessing of the Dataset

Preprocessing is an important step in machine learning, which involves preparing the data for analysis by transforming it into a format suitable for the algorithm. To ensure the optimal performance of ml and dl models for the Microorganism Classification Dataset, it is crucial to follow a systematic approach in preparing the data. This process typically involves Data Cleaning, Feature Encoding, Feature Selection, and Data Splitting.

❖ Data Cleaning

The Microorganism Classification Dataset is an extensive collection of microscopic images designed to aid in the development of ml and dl models for accurate identification and classification of various microorganisms. This dataset comprises eight distinct categories, including Amoeba (75 images), Euglena (170 images), Hydra (75 images), Paramecium (155 images), Rod bacteria (85 images), Spherical bacteria (85 images), Spiral bacteria (75 images), and Yeast (75 images). Each category is organized into separate folders for easy access and a structured dataset.

Before Utilizing this dataset for model development, it is essential to perform data cleaning to ensure the highest quality and consistency. Data cleaning steps for this dataset might include:

- **Removing Duplicate Images:** Examine the dataset for any duplicate images, as they can introduce biases and negatively impact model performance. Eliminate duplicates to maintain a balanced dataset.
- **Image Quality Control:** Inspect the images to ensure they are of high quality and free from artifacts or noise that could impair model performance. Discard any low-quality images or enhance them using image processing techniques if possible.
- **Image Resizing and Normalization:** Standardize image dimensions and scale the intensity values to a common range, such as [0, 1] or [0, 255]. This step ensures that the input data is consistent, enabling more efficient model training.
- **Augmentation:** In cases where the dataset has a limited number of images for certain categories, consider using data augmentation techniques like rotation, flipping, or scaling to artificially increase the number of images and improve the model's ability to generalize.
- **Label Verification:** Verify that the labels assigned to each image correctly correspond to the microorganism

type. Correct any mislabeled images to prevent inaccuracies in the model's training process.

- **Splitting the Dataset:** Divide the dataset into training, validation, and testing subsets to enable model evaluation and avoid overfitting. Ensure the distribution of microorganism types is consistent across all subsets.

By carefully cleaning the Microorganism Classification Dataset, researchers and practitioners can establish a strong foundation for developing advanced machine learning and deep learning models that accurately and efficiently classify microorganisms, ultimately contributing to the broader goals of microbiological research and its applications.

❖ Feature Encoding

Feature Encoding, within the realm of image classification, involves deriving significant characteristics from the images that serve as input for machine learning or deep learning models. This process can be executed through conventional image processing methods like edge detection, texture evaluation, and color histogram analysis, or by employing convolutional neural networks (CNNs) to autonomously discern the most pertinent features present in the images.

❖ Feature Selection

To decrease the model's dimensionality and computing complexity, a procedure known as feature selection is used to extract the most crucial features from encoded data. This process can lessen the risk of overfitting and enhance the readability of the model. To choose the most informative features for the classification task, strategies like as linear discriminant analysis (LDA), principal component analysis (PCA), or recursive feature elimination (RFE) might be used.

❖ Data Splitting

The dataset should be divided into distinct subsets for training, validation, and testing in order to evaluate the model's performance and avoid overfitting. The training set is used to tune the model's parameters and select the best-performing model, while the validation set is used to tune hyperparameters and select the best-performing model, and the testing set provides an unbiased estimate of the model's performance on unobserved data. It is essential to maintain a consistent distribution of microorganism types across all subsets to ensure a fair evaluation of the model's capabilities. By following these steps in preparing the Microorganism Classification Dataset, researchers and practitioners can develop more accurate and efficient machine learning and deep learning models for classifying microorganisms, ultimately contributing to the advancement of microbiological research and its applications in various fields.

3.2 Data Analysis

The given dataset is organized into eight different folders based on microorganism type, including Euglena, Amoeba, Hydra, Rod bacteria, Spherical bacteria, Spiral bacteria, Paramecium and Yeas are all types of bacteria. This organization allows for efficient retrieval and analysis of images or data related to specific microorganisms, making it useful for researchers studying microorganisms or for machine learning and deep learning approaches to microorganism image recognition. The pie chart represents the class distribution of microorganisms and provides insight into the prevalence of different microorganisms. According to the chart, Euglena is the most prevalent microorganism, making up 21.3% of the total distribution. Paramecium and rod bacteria follow closely, accounting for 10.8% and 10.9% respectively. The distribution of spherical bacteria, spiral bacteria, and yeast is relatively consistent, with each comprising 9.51% of the total distribution. Hydra, on the other hand, is the least prevalent microorganism, accounting for only 9.63% of the total distribution. This pie chart can be useful for researchers and scientists who are interested in studying specific microorganisms or analyzing the prevalence of different types of microorganisms in their research or experiments. [Fig2].

The Bar Plot displays the class distribution of microorganisms and provides a quick reference for the prevalence of each type. The plot shows that Euglena is the most common microorganism in the dataset, with 170 occurrences, followed by Paramecium with 155. The remaining microorganisms are more evenly distributed, with Amoeba, Hydra, Spherical Bacteria, Spiral Bacteria, and Yeast all appearing 75 times, and Rod Bacteria appearing 85 times. This Bar Plot can be useful for researchers studying specific microorganisms or for identifying the prevalence of different types of microorganisms in a dataset. [Fig3].

A heatmap can be a useful tool for microorganism image recognition. Heatmaps can be used to visualize the prevalence of different types of microorganisms in a given dataset or image. The heatmap can be generated based on the density of microorganisms in a particular area, allowing researchers to quickly identify regions with high or low levels of specific microorganisms. Heatmaps can also be used to represent the similarity or dissimilarity between different microorganisms based on their features or characteristics. This could be useful for identifying patterns or relationships between different types of microorganisms and aiding in the classification or identification of specific microorganisms. Overall, heatmaps can provide valuable insights into the distribution, similarity, or dissimilarity of different microorganisms in an image or dataset. They can be a useful tool for researchers studying microorganisms or for ML and DL approaches to microorganism image recognition. [Fig4]

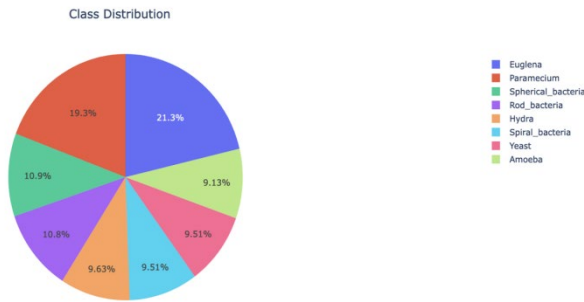


Fig 2. Percentage of microorganisms



Fig 3. Bar Plot microorganisms

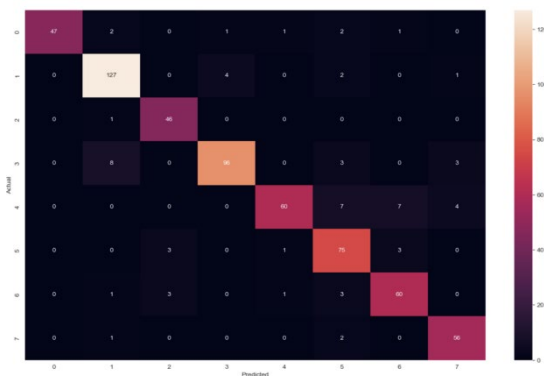


Fig 4. Heatmap

4 Experimental Analysis

Conducting an experimental analysis is vital for determining the effectiveness of ml and dl models tailored for the Microorganism Classification Dataset. Through a series of steps, researchers can evaluate the model's capabilities, recognize potential improvements, and make well-informed choices for ongoing optimization.

4.1 Confusion Matrix

A confusion matrix is employed to assess the performance of the image recognition model. The model's training results demonstrate a loss of 0.3296 and an accuracy of 0.9020, while the validation results show a loss of 1.0177 and an accuracy of 0.7200. This confusion matrix effectively captures the model's performance in categorizing microorganisms, providing valuable insights into the areas where the model excels or requires improvement, ultimately supporting the advancement of intelligent image recognition techniques in the field of microbiology. The confusion matrix is used to assess the image recognition model's performance in classifying microorganisms. The matrix provides a clear representation of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

True positives (TP) represent the instances where the model accurately identifies the presence of a specific microorganism. True negatives (TN) are instances where the model correctly recognizes the absence of that microorganism. False positives (FP) occur when the model incorrectly predicts the presence of a microorganism, while false negatives (FN) are instances when the model fails to identify the presence of a microorganism when it is indeed present.

By analyzing these values, the confusion matrix allows researchers to evaluate the model's precision, recall, and overall accuracy. Precision is calculated as $TP / (TP + FP)$, which measures the model's ability to correctly identify the presence of a microorganism. Recall, calculated as $TP / (TP + FN)$, assesses the model's ability to recognize all instances of a specific microorganism. Ultimately, these metrics contribute to understanding the model's performance and identifying areas for improvement in intelligent image recognition techniques within microbiology.

4.2 Performance of the Three CNN Model

A number of metrics, such as validation loss, validation accuracy, precision, recall, and accuracy, are used in ml to assess a model's performance. These indicators are essential

for assessing the model's accuracy and efficacy. The model's ability to generalise to new data is measured by the validation loss. It measures the discrepancy between the validation dataset's actual output and the output that was predicted. A smaller validation loss denotes a model with stronger generalisation capabilities. The percentage of correctly categorised cases in the validation dataset is referred to as validation accuracy. It is an indicator of how well a model can classify fresh, unstudied data. The ratio of actual positive instances to all the positive instances the model predicts as positive is known as precision.

The performance of the three CNN models varied significantly. The first model achieved an accuracy of 92.71% with a precision of 93.89% and a recall of 91.65%. The validation loss was 0.2394 with a validation accuracy of 92.71%, a validation precision of 93.82%, and a validation recall of 91.44%.

The second model had an accuracy of 90.20% and a validation loss of 1.0177, with a validation accuracy of 72.00% [Fig. 5] For the third model I had an accuracy of 90.20% 0.522%. Overall, the first CNN model performed the best with the highest accuracy, precision, and recall. However, it is essential to note that the performance of the CNN models can vary based on various factors, such as the dataset used, the number of layers, and the training parameters.

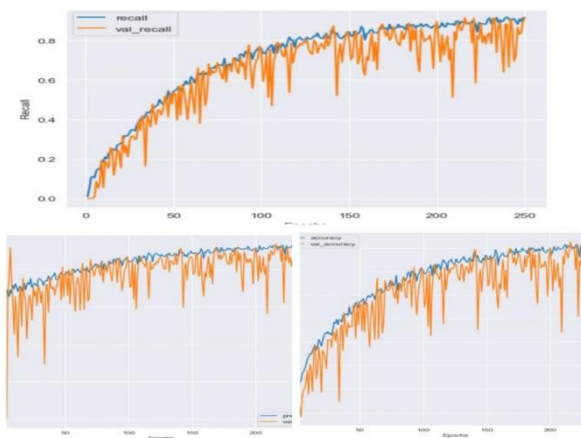


Fig 5. Performance of the three CNN model

5 Result

The results of this study suggest that dl algorithms, particularly Convolutional Neural Networks (CNN), are highly effective for microorganism image recognition tasks. Three different CNN models were tested in this study, resulting in test accuracies of 0.52, 0.90, and 0.93, demonstrating the superior accuracy of deep learning models in image classification tasks. In contrast, traditional machine

learning algorithms were found to be less effective for microorganism image recognition tasks in this study, with lower accuracy rates observed. This further emphasizes the importance of using appropriate techniques and models for specific tasks in order to achieve the most accurate results. The study highlights the utility of deep learning algorithms, specifically CNN models, for microorganism image recognition tasks. These findings are in line with previous research, which has shown that deep learning algorithms can achieve higher accuracy rates in image classification tasks compared to traditional machine learning algorithms. Overall, the results of this study demonstrate the importance of utilizing appropriate techniques and models for specific tasks, and highlight the potential for deep learning algorithms, particularly CNN models, in advancing the field of microorganism image recognition.

6 Conclusion

The study of microorganisms through intelligent image recognition is a complex and challenging task. However, advancements in ML and DL algorithms, specifically CNN models, have shown great promise in improving microorganism image recognition accuracy. The use of deep learning algorithms in microorganism image recognition tasks has demonstrated superior accuracy rates compared to traditional machine learning algorithms. This highlights the importance of utilizing appropriate techniques and models for specific tasks in order to achieve the most accurate results. Despite the promising results of this study, there are still challenges that need to be addressed in microorganism image recognition. For example, the quality and quantity of data available can greatly impact the accuracy of models. Additionally, microorganisms can be highly variable in their characteristics, making it difficult to identify and classify them accurately. Further advancements in ml and dl algorithms are necessary to overcome these challenges and improve accuracy rates in microorganism image recognition. The use of techniques such as data augmentation and transfer learning can potentially improve the accuracy of models by allowing them to learn from a larger and more diverse set of data. In summary, the study of microorganisms through intelligent image recognition is a rapidly evolving field with the potential to drive significant advancements in various areas, including clinical microbiology, agriculture, medical science, and food production. The use of machine learning and deep learning algorithms, specifically CNN models, has demonstrated great promise in improving accuracy rates in microorganism image recognition, and further advancements in these techniques can potentially lead to even more significant improvements in accuracy and effectiveness.

References

1. Environmental Research; Study Results from D.H. Ma and Colleagues Broaden Understanding of Environmental

- Research (Biological removal of antiandrogenic activity in gray wastewater and coking wastewater by membrane reactor process) (p. 880). (2015). NewsRx.
- Lin, F., Duan, Q.-Y., & Wu, F.-G. (2020). Conjugated Polymer-Based Photothermal Therapy for Killing Microorganisms. *ACS Applied Polymer Materials*, 2(10), 4331–4344. <https://doi.org/10.1021/acsapm.0c00718>
 - Gracias, K. S., & McKillip, J. L. (2011). Triplex PCR-based detection of enterotoxigenic *Bacillus cereus* ATCC 14579 in nonfat dry milk. *Journal of Basic Microbiology*, 51(2), 147–152. <https://doi.org/10.1002/jobm.200900348>
 - Daims, H., & Wagner, M. (2018). Nitrospira. *Trends in Microbiology* (Regular Ed.), 26(5), 462–463. <https://doi.org/10.1016/j.tim.2018.02.001>
 - Barbedo, J. G. A. (2012). Method for Counting Microorganisms and Colonies in Microscopic Images. 2012 12th International Conference on Computational Science and Its Applications, 83–87. <https://doi.org/10.1109/ICCSA.2012.23>
 - Dazzo, F., Sexton, R., Jain, A., Makhoul, A., Shears, M., Gusfa, D., Handelsman, S., Niccum, B., & Onsay, D. (2017). Influence of Substratum Hydrophobicity on the Geomicrobiology of River Biofilm Architecture and Ecology Analyzed by CMEIAS Bioimage Informatics. *Geosciences* (Basel), 7(3), 56. <https://doi.org/10.3390/geosciences7030056>
 - Antharam, V. C., McEwen, D. C., Garrett, T. J., Dossey, A. T., Li, E. C., Kozlov, A. N., Mesbah, Z., & Wang, G. P. (2016). An Integrated Metabolomic and Microbiome Analysis Identified Specific Gut Microbiota Associated with Fecal Cholesterol and Coprostanol in *Clostridium difficile* Infection. *PloS One*, 11(2), e0148824–e0148824. <https://doi.org/10.1371/journal.pone.0148824>
 - Puchkov, E. O. (2019). Quantitative Methods for Single-Cell Analysis of Microorganisms. *Microbiology* (New York), 88(1), 1–14. <https://doi.org/10.1134/S0026261719010120>
 - Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H.-Q., & Fredrickson, J. K. (2016). Extracellular electron transfer mechanisms between microorganisms and minerals. *Nature Reviews. Microbiology*, 14(10), 651–662. <https://doi.org/10.1038/nrmicro.2016.93>
 - Fan, W., Huang, X., Liu, K., Xu, Y., & Chi, Z. (2023). Advanced upcycling of agro-industrial co-products of corn via different microorganisms. *Biomass & Bioenergy*, 168, 106669. <https://doi.org/10.1016/j.biombioe.2022.106669>
 - Rani, P., Kotwal, S., Manhas, J., Sharma, V., & Sharma, S. (2022). Machine Learning and Deep Learning Based Computational Approaches in Automatic Microorganisms Image Recognition: Methodologies, Challenges, and Developments. *Archives of Computational Methods in Engineering*, 29(3), 1801–1837. <https://doi.org/10.1007/s11831-021-09639-x>
 - Zhang, Y., Jiang, H., Ye, T., & Juhas, M. (2021). Deep Learning for Imaging and Detection of Microorganisms. *Trends in Microbiology* (Regular Ed.), 29(7), 569–572. <https://doi.org/10.1016/j.tim.2021.01.006>
 - Liang, C.-M., Lai, C.-C., Wang, S.-H., & Lin, Y.-H. (2021). Environmental microorganism classification using optimized deep learning model. *Environmental Science and Pollution Research International*, 28(24), 31920–31932. <https://doi.org/10.1007/s11356-021-13010-9>
 - Kulwa, F., Li, C., Zhang, J., Shirahama, K., Kosov, S., Zhao, X., Jiang, T., & Grzegorzec, M. (2022). A new pairwise deep learning feature for environmental microorganism image analysis. *Environmental Science and Pollution Research International*, 29(34), 51909–51926. <https://doi.org/10.1007/s11356-022-18849-0>
 - Yang, H., Zhao, X., Jiang, T., Zhang, J., Zhao, P., Chen, A., Grzegorzec, M., Qi, S., Teng, Y., & Li, C. (2022). Comparative Study for Patch-Level and Pixel-Level Segmentation of Deep Learning Methods on Transparent Images of Environmental Microorganisms: From Convolutional Neural Networks to Visual Transformers. *Applied Sciences*, 12(18), 9321. <https://doi.org/10.3390/app12189321>
 - Zhang, J., Li, C., Rahaman, M. M., Yao, Y., Ma, P., Zhang, J., Zhao, X., Jiang, T., & Grzegorzec, M. (2022). A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches. *The Artificial Intelligence Review*, 55(4), 2875–2944. <https://doi.org/10.1007/s10462-021-10082-4>
 - Ma, P., Li, C., Rahaman, M. M., Yao, Y., Zhang, J., Zou, S., Zhao, X., & Grzegorzec, M. (2023). A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches. *The Artificial Intelligence Review*, 56(2), 1627–1698. <https://doi.org/10.1007/s10462-022-10209-1>
 - Biaisoni, R., Di Marco, E., Squillario, M., Barla, A., Piccolo, G., Ugolotti, E., Gatti, C., Minuto, N., Patti, G., Maghnie, M., & d'Annunzio, G. (2020). Gut Microbiota in T1DM-Onset Pediatric Patients: Machine-Learning Algorithms to Classify Microorganisms as Disease Linked. *The Journal of Clinical Endocrinology and Metabolism*, 105(9), E3114–E3126. <https://doi.org/10.1210/clinem/dgaa407>
 - Nocedo-Mena, D., Cornelio, C., Camacho-Corona, M. del R., Garza-Gonzalez, E., Waksman de Torres, N., Arrasate, S., Sotomayor, N., Lete, E., & Gonzalez-Diaz, H. (2019). Modeling Antibacterial Activity with Machine Learning and Fusion of Chemical Structure Information with Microorganism Metabolic Networks. *Journal of Chemical Information and Modeling*, 59(3), 1109–1120. <https://doi.org/10.1021/acs.jcim.9b00034>
 - Pei, Z., Liu, S., Jing, Z., Zhang, Y., Wang, J., Liu, J., Wang, Y., Guo, W., Li, Y., Feng, L., Zhou, H., Li, G., Han, Y., Liu, D., & Pan, J. (2022). Understanding of the interrelationship between methane production and microorganisms in high-

- solid anaerobic co-digestion using microbial analysis and machine learning. *Journal of Cleaner Production*, 373, 133848. <https://doi.org/10.1016/j.jclepro.2022.133848>
21. Huang, T.-S., Lee, S. S.-J., Lee, C.-C., & Chang, F.-C. (2020). Detection of carbapenem-resistant *Klebsiella pneumoniae* on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using supervised machine learning approach. *PloS One*, 15(2), e0228459–e0228459. <https://doi.org/10.1371/journal.pone.0228459>
 22. Nakanishi, A., Fukunishi, H., Matsumoto, R., & Eguchi, F. (2022). Development of a Prediction Method of Cell Density in Autotrophic/Heterotrophic Microorganism Mixtures by Machine Learning Using Absorbance Spectrum Data. *Biotech (Basel)*, 11(4), 46. <https://doi.org/10.3390/biotech11040046>
 23. Yu, T., Su, S., Hu, J., Zhang, J., & Xianyu, Y. (2022). A New Strategy for Microbial Taxonomic Identification through Micro-Biosynthetic Gold Nanoparticles and Machine Learning. *Advanced Materials (Weinheim)*, 34(11), e2109365–n/a. <https://doi.org/10.1002/adma.202109365>
 24. Maeda, Y., Kohketsu, H., Tanaka, Y., Sugiyama, Y., Kogiso, A., Lim, T.-K., Harada, M., Yoshino, T., Matsunaga, T., & Tanaka, T. (2020). (Invited) Rapid and Robust Discrimination of Food-Contaminating Microorganisms Guided By Machine Learning. *Meeting Abstracts (Electrochemical Society)*, MA2020-02(44), 2812–2812. <https://doi.org/10.1149/MA2020-02442812mtgabs>
 25. Zhang, J., & Fernando, S. D. (2023). Identification of Fungicide Combinations Targeting *Plasmopara viticola* and *Botrytis cinerea* Fungicide Resistance Using Machine Learning. *Microorganisms (Basel)*, 11(5), 1341. <https://doi.org/10.3390/microorganisms11051341>
 26. Gado, J. E., Beckham, G. T., & Payne, C. M. (2020). Improving Enzyme Optimum Temperature Prediction with Resampling Strategies and Ensemble Learning. *Journal of Chemical Information and Modeling*, 60(8), 4098–4107. <https://doi.org/10.1021/acs.jcim.0c00489>
 27. Truong, V. K., Chapman, J., & Cozzolino, D. (2021). Monitoring the Bacterial Response to Antibiotic and Time Growth Using Near-infrared Spectroscopy Combined with Machine Learning. *Food Analytical Methods*, 14(7), 1394–1401. <https://doi.org/10.1007/s12161-021-01994-6>
 28. Bemani, A., Kazemi, A., & Ahmadi, M. (2023). An insight into the microorganism growth prediction by means of machine learning approaches. *Journal of Petroleum Science & Engineering*, 220, 111162. <https://doi.org/10.1016/j.petrol.2022.111162>
 29. Abdullah, A. A., Aziz, A. N. A., Kanaya, S., & Ranjan Dash, S. (2019). Classification of Microorganism Species Based on Volatile Metabolite Contents Similarity. *Journal of Physics. Conference Series*, 1372(1), 12061. <https://doi.org/10.1088/1742-6596/1372/1/012061>
 30. Helena Tavares Kennedy. (2021). Biofuels Digest: Pollution pods, machine learning, fermentation protein-producing microorganisms, Microsoft's look at algae and fungi for green data centers and more: The Digest's Top 8 Innovations for the week of November 4th. In *Newstex Trade & Industry Blogs*. Newstex.
 31. Ensemble Learning Framework with GLCM Texture Extraction for Early Detection of Lung Cancer on CT Images, Sara A Althubiti, Sanchita Paul, Rajnikanta Mohanty, Sachi Nandan Mohanty, Fayadh Alenezi, Kemal Polat, *Computational and Mathematical Methods in Medicine (Hindawi)*, 2022, doi.org/10.1155/2022/2733965
 32. A Novel Approach for Diabetic Retinopathy Screening Using Asymmetric Deep Learning Feature, Pradeep Kumar Jena, Bonomali Khuntia, Charulata Palai, Manjushree Nayak, Tapas Kumar Mishra, Sachi Nandan Mohanty, *Big Data Cognitive Computing (2023)*, Vol 7, Issue 1, 25, <https://doi.org/10.3390/bdcc7010025>, ISSN: 2504-2289