MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

U-Net Guided EfficientNet Framework for Accurate Bacterial Colony Counting and Species Recognition

M. Sivapriya[1], N. Senthilkumaran[2*]
[1, 2] Department of Computer Science and Applications,
[1, 2] The Gandhigram Rural Institute (Deemed to be University), Gandhigram,
Dindigul – 624 302, Tamilnadu, India.
muthuthiru17@gmail.com

Abstract

Accurate enumeration of bacterial colonies on agar plates is a critical step in microbiology, enabling reliable assessment of microbial growth, contamination, and antimicrobial activity. Traditional manual counting is labor intensive, subjective, and prone to error, particularly when colonies overlap or vary in size and contrast. To address these challenges, we propose a two stage deep learning framework that integrates U-Net based image segmentation with EfficientNet family regression and classification models for automated colony detection, counting, and species identification. In this first stage, a U-Net architecture performs pixel level segmentation of agar plate images to isolate colony regions from the background. This preprocessing step effectively handles variations in illumination and background noise, providing high quality binary masks that delineate individual colonies. In the second stage, segmented images are processed by a series of EfficientNet variants (B0 through B7) to simultaneously estimate total colony count and classify colonies by species. Training and evaluation were performed on a curate dataset of annotated agar plate images spanning five clinically relevant species Bacillus subtilis (B.subtilis), Candida albicans (C.albicans), Escherichia coli (E.coli), and Pseudomonas aeruginosa (P.aeruginosa), Staphylococcus aureus (S.aureus) with ground truth counts and species labels provided in JSON format. Comprehensive experiments demonstrate that all EfficientNet models achieve strong predictive accuracy with EfficientNetB7 delivering the best overall performance Mean Absolute Error (MAE=2.60), Root Mean Square Error (RMSE=12.57), Coefficient of Determination (R^2=0.67), accuracy=96.6%. Per-species F1-Scores range from 0.93 to 0.96 indicating consistent detection across diverse colony morphologies. The proposed U-Net + EfficientNet pipeline offers a fast, reproducible, and scalable solution for high throughput microbial analysis, significantly reducing manual effort while improving the reliability of colony enumeration and species classification in clinical and research laboratories.

Keywords: Microorganisms, Microbial Images, Segmentation, Regression, Classification, Counting.

I. INTRODUCTION

Microorganisms are small organisms that cannot be seen with the naked eye but can be seen with an electron or light microscope. There are many distinct types of microbes, and their classification standards vary. Microorganisms consist primarily of bacteria, viruses, fungus, and some algae. Bacteria are single-cell organisms with small sizes, simple structures, no nuclei, membrane organelles, and cytoskeletons. It is widely distributed in soil and water, and the majority of it is decomposed by bacteria at the bottom of the biological chain, such as Escherichia coli. Some bacteria function as both consumers and producers. Producers include sulfur bacteria and iron bacteria [1]. They can use inorganic elements to create the organic compounds they require. Rhizobia can devour organic compounds produced by the photosynthesis of legumes. The virus is a bacterium that may propagate and infect other species. It has a basic structure and is small. It comprises only one form of nucleic acid, namely deoxyribonucleic acid (DNA) virus and ribonucleic acid (RNA) virus. It must parasitize live cells and reproduce via replication. Fungus is eukaryotic microorganisms that may create spores through both asexual and sexual reproduction, such as mold, yeast, and mushrooms. Globally prevalent, tinea pedis is a type of foot skin disease brought on by pathogenic fungus. Since human feet and toes lack sebaceous glands, an environment deficient in fatty acids and with inadequate air circulation is favorable for the development of filamentous fungus. Algae are eukaryotic protozoa, with the majority of them being aquatic organisms capable of photosynthesis [2].

Some microbes are hazardous to humans because they cause food breakdown, infection, and sickness, while others are beneficial. Penicillin was a groundbreaking medicinal breakthrough that saved countless lives. Yeast is commonly employed in industrial fermentation, ethanol generation, and human food production. Certain microbes have the ability to break down plastics, treat gas and wastewater, and have enormous promise for renewable resources. Humans may break down and absorb food and harmful substances with the aid of the many bacteria found in healthy human intestines. Some microbes have harmful effects on the human body and industrial processes. Microorganisms are crucial to human existence and production [3]. Thus, good microbes should be widely employed, while dangerous microorganisms should be avoided.

Bacterial growth is commonly measured in microbial studies using colony formation on solid agar plates. Colonies emerge as observable clusters of cells after incubation, with each preferably originating from a single bacterial cell or a small group [4]. The number of colonies is thus proportional to the number of viable organisms in the initial sample and is stated in terms of Colony Forming Unit (CFU). Agar plates are frequently shot with digital cameras to simplify documentation and study, yielding microbiological images that capture colony size, form, and distribution. However, overlapping colonies, varied growth patterns, and fluctuating illumination are common issues in these images, making colony quantification difficult.

Traditionally, professional microbiologists visually count colonies on agar plates to enumerate them manually. While simple, this procedure is time consuming, labor-intensive, and subject to human error. The problem is compounded when dealing with plates containing hundreds of colonies, unevenly shaped colonies, or faint growth at the edges of plates. To reduce manual labor, semi-automated colony counters using image thresholding, Hough Transform algorithms have been created. However, these methods frequently exhibit poor adaptability to varied microbial [5] species and imaging circumstances, resulting in low reliability for high-throughput microbiological applications.

Deep learning (DL) has provided new opportunities for automated bacterial colony counts with greater accuracy and scalability. Convolutional Neural Network (CNN) [6] has demonstrated extraordinary performance in biomedical imaging tasks such as segmentation, classification [7], object detection [8], and counting [9]. By learning complex spatial and morphological information directly from microbiological images. CNN overcome many of the limitations of manual and rule-based approaches. For example, DL models can effectively handle overlapping colonies, irregular shapes, and variations in lighting conditions. Several studies have already demonstrated the potential of DL frameworks such as U-Net [10], YOLO [11], ResNet, Faster R-CNN, and Mask R-CNN[12] for microbial colony detection and counting, achieving significantly higher accuracy than conventional approaches.

Among the various DL architectures, the EfficientNet [13] family stands out for its novel compound scaling strategy, which balances network depth, width, and resolution to maximize performance. EfficientNet models (B0-B7) achieve state-of-the-art accuracy while using fewer parameters and computational power than typical CNN. This efficiency makes them ideal for biomedical image analysis [14], where accuracy and resource constraints are crucial. In the domain of bacterial colony counting, EfficientNet models offer a scalable framework that can accommodate microbial images of different complexity. Smaller forms, such as EfficientNet-B0 and B1 are lightweight and ideal

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

for real time applications, whereas larger variants, such as B6 and B7, improve accuracy for complicated and high resolution microbiological images.

This study aims to develop and evaluate an automated bacterial colony counting framework based on the EfficientNet family of models. We thoroughly evaluate the performance of EfficientNet versions (B0-B7) on microbiological images with annotated colony counts, focusing on robustness, accuracy, and computational efficiency. The suggested method overcomes the constraints of manual and semi-automated approaches by offering a scalable, reliable, and high-throughput solution for colony enumeration in microbiological research, clinical diagnostics, and industrial applications.

The rest of the paper is structured as follows. In Section II review of related works, In Section III suggested methodology U-Net and EfficientNet (B0-B7) models. In Section IV results and discussion. Finally, in Section V conclusions were given.

II. LITERATURE REVIEW

This section outlines image analysis strategies for enumerating bacterial colonies, emphasizing accuracy and automation. Researchers use these methodologies in a variety of fields, including food and water quality assessment, clinical microbiology, environmental monitoring, and pharmaceutical testing. Early image processing algorithms for isolating colonies from agar [15] backgrounds included edge detection, global and adaptive thresholding, watershed transformation, morphological operations. Although these traditional methods are effective for detecting microorganisms, they are extremely sensitive to lighting changes and overlapping colonies, allowing potential for improvement in segmentation robustness along with complete automation. Expanding datasets to include a wider range of bacterial strains and plate conditions has been demonstrated to improve the dependability of these methods and encourage wider implementation in research and industry facilities.

Microbial colony counting can be done either manually or automatically. Manual enumeration is fundamentally subjective and prone to inconsistency; even the same observer may report different counts while studying the same plate multiple times. Automated counting based on image analysis and machine learning provides objective, repeatable findings while drastically reducing labor. These procedures usually include differentiating colonies from the agar background and quantifying the discovered regions of interest.

Recent DL developments have improved automatic counting even more. CNN may learn hierarchical features directly from raw images, exceeding hand crafted approaches in terms of accuracy and robustness. U-Net [10], Faster R-CNN, and Mask R-CNN, ResNet [16], YOLO architectures have been widely used for microbial image analysis, with excellent results in both segmentation and instance detection. The EfficientNet family continues this trend by introducing a compound scaling technique that balances network depth, width, and input resolution to achieve high accuracy with fewer parameters. Variants ranging from B0 to B7 provide a variety of models that can be matched to varied computing budgets while capturing fine colony morphology. This review focuses on DL approaches, specifically EfficientNet, as a foundation for precise and scalable bacterial colony enumeration.

Khan et al., [17] offered an automated method for measuring the number of bacterial cells in water samples. To identify E. coli germs, a CNN was developed. The CNN model has six convolutional layers following two fully linked ones. The R-CNN deep learning model was used to automatically count bacterial cells within a colony. The models use previously trained ResNet-50, resulting in the use of TL. The ResNet classification layers were replaced by nine neurons representing nine R-CNN classes. The proposed models achieve 97% accuracy in recognizing and counting E-coli bacteria in colonies.

Rahman et al.,[18] attempted to determine the structural properties of G20 bacteria, which induce corrosion on steel surfaces. SEM images are used as data samples. We employed Deep Convolutional Neural Network (DCNN) and Mask Region CNN (RCNN) to segment bacterial instances and identify clustered bacteria. To develop deep learning architectures, the authors employ a previously created platform called DeepLABv3+.A pre-trained architecture called RCNN is used to extract feature maps from images. Experienced human specialists' deductions were compared. Results show that RCNN and DCNN outperform manual and traditional methods for detecting bacteria in biofilms, with an accuracy of 81%.

In [19], Alhammad et al., provided a CNN-based method for classifying entire slide images of three gram-positive bacteria species. Bacteria were separated from the background using pre-processed data. Bacterial pictures were split and sorted into three categories using a classifier. Bacteria were separated from the background using a pre-trained ResNet. Finally found that the architecture of ResNet can distinguish between three gram-positive bacteria with 81% accuracy.

Shwetha et al.,[20] suggest a deep CNN-based technique for detecting bacilli. SqueezeNet, VGG-16 and ResNet trained with tested for determine the best models for identifying bacilli. The models may have employed TL based on the use of pre-trained architecture. The colorization and K-means clustering techniques were used to distinguish the bacterial images from the backdrop. Images were inserted into the model after being resized to 224x224x3. Between layers, sigmoid activation functions and ReLU were used.

The Casado et al.,[21] employed image segmentation to identify bacteria-infested areas in surrounding images and use them as input for the model. Bacterial segments were obtained using image processing techniques like noise reduction and erosion. CNNs such as FBNet, ResNet50, EfficientNet B3, and ResNet101 were utilized for fine tuning and TL. Following additional fine-tuning, a linear layer was used in place of the final convolutional layer. Input images were used to train every model.

In [22], Kakishita et al., proposed a method to segment and classify 6 species of bacteria. Then, SEM images were divided into backgrounds and foregrounds using the UNet architecture. Prior to counting the number of cells, the species of bacteria were categorized using VGG16. Then separated classification and segmentation to allow for autonomous learning and increased accuracy. The proposed technique outperformed traditional CNN, with a classification accuracy of 95.8%.

In [23], Sundar et al., attempt to identify dangerous microbes that can cause major health problems by contaminating the bloodstreams. Then used U-Net, to detect whether bacterial cells are present in images captured by a dark field microscope. The authors highlight U-Net model effectiveness in object detection and image segmentation. Then identified four models. The model of first utilized early pausing to prevent overfitting if training set performance didn't improve after few cycles. Early stop is not used in the second model. Next, model introduces function of loss. Finally proposed UNet model achieved 96.6% accuracy, according to the data.

Abeyrathna et al., [24], intend to measuring quantity and bacterial cell length, as well as biofilms covered total area. The authors assert that they have resolved the issue of segmentation in cells that overlap or are close to one another. The authors used UNet models and bacterial cell occurrences in biofilms by segmenting, counting, and measuring them using a "region-based ellipse fitting technique. The UNet encodes and decodes input pictures through convolution and deconvolution. The model generates a logical mask of bacterium cell instances, categorizing each pixel into a certain classes. Next the ellipse approach was used to calculate and distances and centroids, then iteratively assemble similar angles ellipses. Finally this approach achieved a high recall rate of 93.6%.

Khan et al.,[25] proposed an automated approach for detecting and counting E. coli colonies to replace time-consuming, error-prone manual procedures that take 24-48 hours and skilled microscopy. Researchers collected 1,301 groundwater samples from eight districts in Rajasthan, India, and used artificial intelligence tools to conduct quick analysis. A convolutional neural network (CNN) was created to

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

automatically identify E. coli colonies, together with a smartphone app enabling on-site detection. To automate enumeration, a Faster R-CNN model with a graphical user interface (GUI) was used to count colony-forming units on agar plates, yielding 97% accuracy with a 0.10 loss. Model performance was evaluated using F-score, precision, sensitivity, and accuracy, confirming that Faster R-CNN is a dependable and efficient solution for real-time monitoring of water contamination hotspots.

Huang et al., [26] offered an EfficientNet based rock image classification model that incorporates a triplet attention mechanism to overcome the constraints of classic manual and deep learning methods. While EfficientNet creates an efficient network topology using NAS technology and compound scaling, the triplet attention module increases feature representation by capturing both channel and spatial information, increasing classification accuracy. Transfer learning is used using pre-trained model parameters to accelerate convergence and shorten training time. When tested on rock picture datasets, the model achieves 92.6% training accuracy and 93.2% Top-1 test accuracy, exceeding mainstream alternatives while displaying good durability and adaptability for automated geological survey applications.

Traditional methods such as thresholding, morphological operations, and watershed segmentation are sensitive to lighting, colony overlap, and agar texture. Classical ML methods (e.g., SVM, Random Forest) require hand-crafted features and fail under non-uniform illumination. Deep CNN-based single-stage approaches (e.g., ResNet, YOLO) perform poorly when colonies are confluent or transparent.

To overcome these issues, this study proposes a two-stage hybrid pipeline combining U-Net-based segmentation for region isolation with EfficientNet regression and classification for robust colony quantification and species identification. Unlike prior works, the proposed approach integrates pixel-level segmentation with count-level and species-level analysis in a unified workflow, validated on a diverse agar dataset.

III. METHODOLOGY

This section details the entire procedure used to create and develop an EfficientNet based framework for automatic bacterial colony enumeration. The methodology is organized into five stages such as dataset preparation, preprocessing [27] and augmentation, model architecture, training architecture, and evaluation.

3.1 Dataset Preparation

The experiments utilized the AGAR dataset [28], which contains high resolution images of bacterial growth on agar plates. The dataset was curated to reflect real laboratory conditions and was divided into three categories such as countable, uncountable, and empty.

- Countable: plates with distinct, non-overlapping colonies that permit precise enumeration.
- Uncountable: plates exhibiting dense or confluent growth where individual colonies cannot be reliably separated.
- Empty: Negative-control plates showing no bacterial growth.

A total of 4178 images with JSON were collected. The dataset includes seven microbiological species or labels such as Bacillus subtilis (B.subtilis), Candida albicans (C.albicans), Environmental Contamination, Plate Defect, Escherichia coli (E.coli), and Pseudomonas aeruginosa (P.aeruginosa), Staphylococcus aureus (S.aureus). For the countable subset, each image was annotated by trained microbiologist to provide the true colony count. The uncountable and empty subsets were assigned categorical labels to support binary classification. The dataset was divided at the plate level into training (80%), validation (10%), and test (10%) sets, ensuring that no images from the same plate appear across splits. Stratified sampling was applied to maintain balance across colony density levels (sparse, moderate, and dense). Cross-validation was also performed to confirm robustness.

3.2 Preprocessing and Data Augmentation

To ensure consistency across images captured under varying laboratory conditions, a standardized preprocessing pipeline was applied.

Resizing: Images were resized to match the input dimensions of each EfficientNet variant (224 x 224 for B0 up to B7).

Normalization: Pixel values were scaled to the [0, 1] range and standardized using ImageNet mean and standard deviation.

Augmentation: To reduce overfitting and improve model generalization, we employed random rotations, brightness and contrast adjustments, and Gaussion noise. Augmentation was performed on the fly during training to produce a diverse set of inputs.

3.3 Model Architecture

The proposed bacterial colony enumeration framework is built upon U-Net segmentation network with the EfficientNet family of CNN, which provides a scalable and computationally efficient backbone for high resolution image analysis. EfficientNet is chosen because it achieves state-of-the- art on large-scale image recognition [29] tasks while maintaining a favorable trade-off between model complexity and inference speed, making it well suited to laboratory settings where both accuracy and throughput are important.

3.3.1 U-Net Segmentation

The first stage of the proposed pipeline uses a U-Net model to precisely split bacterial colonies from agar plate images. U-Net is a fully CNN created primarily for biological image segmentation, making it ideal for differentiating colonies with varying size, shape, and contrast.

3.3.1.1 Encoder (Contracting Path)

- Consists of series of convolutional blocks, each with two successive 3 x 3 convolutions, batch normalization, and ReLU activation.
- Each block is followed by a 2 x 2 max pooling operation with stride 2, which reduces the spatial dimensions while doubling the number of feature channels.
- This path captures increasingly abstract contextual features necessary to distinguish colonies from complex backgrounds such as agar textures lighting variations, and plate edges.

3.3.1.2 Bottleneck Layer

At the networks deepest level, two additional 3 x 3 convolutions with dropout regularization learn dense, high level representations of colony morphology while mitigating overfitting.

3.3.1.3 Decoder (Expanding Path)

- Mirrors the encoder with a sequence of transposed convolutions to upsample feature maps and recover spatial resolution.
- After each upsampling step, the corresponding encoder feature map is concatenated via skip connections to provide fine grained localization cues that may have been lost during pooling.
- Followed by two 3 x 3 convolutions with ReLU activation to refine the upsampled features

3.3.1.4 Output Layer

- The decoder characteristics are converted to single channel probability map by a final 1 x 1 convolution.
- Sigmoid activation converts this map into a pixel wise probability of belonging to the colony class.

The U-Net is trained on pixel level masks generated from AGAR dataset JSON annotations. It output a high fidelity colony mask which serves as a region of interest map for the next stage.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

3.3.2 EfficientNet Scaling Principles

Unlike conventional CNN that scale depth, width, or input resolution independently, EfficientNet introduces a compound scaling method that uniformly scales these three dimensions using a set of fixed coefficients. This strategy results in a family of models ranging from EfficientNet-B0 (lightweight and fast) to EfficientNet-B7 (deep and high capacity).

- **Depth:** Number of layers or repeated blocks, controlling the models ability to capture hierarchical features.
- **Width:** Number of channels in each layer, improving representation of fine grained features.
- **Resolution:** Input image size, allowing finer spatial detail to be captured.

By balancing these factors, EfficientNet provides flexible configurations for different computational budgets without sacrificing accuracy.

3.3.2.1 Backbone and Feature Extraction

Each EfficientNet variant follows a Mobile Inverted Bottleneck Convolution (MBConv) design with depthwise separable convolutions and squeeze-and-excitation (SE) blocks.

- **MBConv Blocks:** use depthwise convolutions to reduce parameter count while retaining spatial information.
- **SE Attention:** Learns channel wise feature recalibration, enabling the network to emphasize biologically relevant regions such as colony textures and boundaries.
- **Swish Activation:** Provides smooth, non-linear transformations that improve gradient flow compared to ReLU.

These architectural elements are advantageous for distinguishing subtle intensity differences between bacterial colonies and agar background, particularly in plates with low contrast or overlapping colonies.

3.3.2.2 Dual-Head Output Design

To handle the heterogeneous AGAR dataset, which includes countable, uncountable, and empty plates. The backbone is extended with a dual-head output layer.

A. Regression Head

- **Global Average Pooling (GAP):** Aggregates spatial feature maps into a single vector, making the model invariant to colonies absolute location on the plate.
- **Fully Connected Dense Layer:** Outputs a single continuous value representing the predicted colony count for plates classified as countable.
- **Activation:** ReLU activation ensures non-negative outputs, consistent with real colony counts.

B. Classification Head

- A parallel branch processes the same pooled features to perform multi-class classification (Countable, Uncountable, and Empty) using softmax activation.
- This branch ensures the system first determines whether numerical enumeration is appropriate before producing a count.

The dual-head design allows end-to-end multitask learning, improving feature sharing and overall robustness while avoiding the need for two separate models.

3.3.2.3 Transfer Learning and Fine-Tuning

Because the AGAR dataset is relatively small compared to ImageNet, the network is initialized with ImageNet pretrained weights to leverage rich generic visual features. Fine tuning is performed in two stages such as

- Freeze the lower MBConv blocks and train only the top dense layers to adapt to microbiological images.
- Gradually unfreeze deeper layers with a low learning rate to refine the backbone for colony textures and agar background patterns.

This strategy speeds convergence and reduces overfitting while ensuring the model learns domain specific representations.

3.3.2.4 Regularization and Optimization

To further enhance generalization

- Dropout Layers (typically 0.3-0.5) are applied after fully connected layers.
- Batch Normalization stabilizes intermediate feature distributions, enabling higher learning rates.
- Weight Decay (L2 Regularization) discourages over-complex models.

The network is trained with the Adam optimizer, using a combined loss function such as Mean Squared Error (MSE) for regression plus Categorical Cross-Entrophy for classification, with task- specific weighting to balance both objectives.

3.3.2.5 Rationale for Architecture

The chosen architecture provides several key benefits for bacterial colony counting

- **Scalability:** EfficientNet-B0 can be deployed on resource constrained devices for rapid screening, while EfficientNet-B7 offers maximum accuracy for high precision laboratory analysis.
- **Fine Grained Feature Sensitivity:** Depth wise separable convolutions and SE attention allow the network to capture small, faint colonies that might be missed by simpler CNN.
- **Integrated Decision Logic:** the dual-head design reduces post processing steps by combining count regression and plate classification within a single unified model.

The final model architecture leverages EfficientNet compound scaling and a lightweight MBConv block, enriched with squeeze-and-excitation attention, and is extended by dual regression and classification heads. This architecture yields a flexible, end-to-end solution capable of handling the diverse visual challenges of bacterial colony enumeration across the AGAR dataset. In Fig.1 displays the proposed pipeline follows a two-stage process:

Stage 1 – Segmentation (U-Net): Input RGB agar plate → binary mask (colony regions).
Stage 2 – Regression & Classification (EfficientNet): Input = segmented mask (3-channel grayscale).
Outputs:

- **Regression head:** Predicts colony count.
- **Classification head:** Predicts bacterial species class.

During inference, U-Net masks are first generated and then fed to EfficientNet for count and species prediction.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

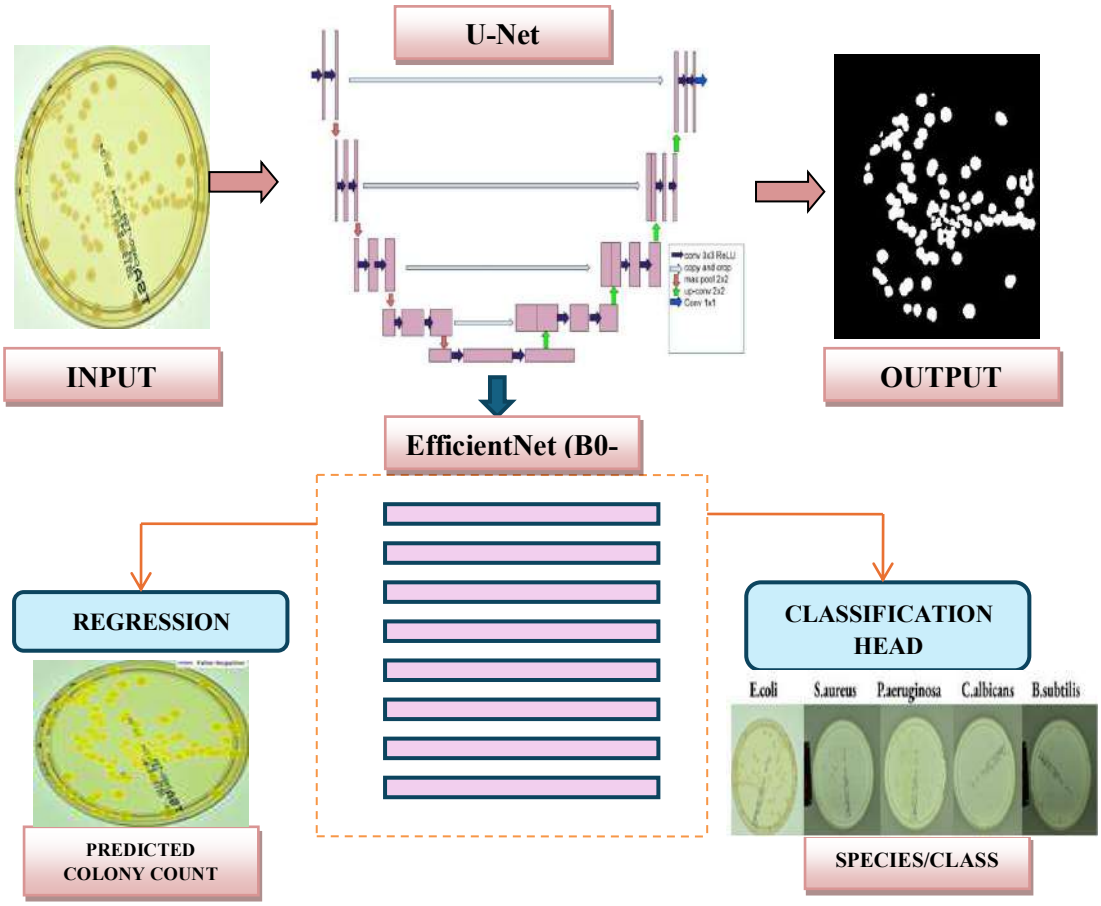

Fig.1. Workflow of Bacterial Colony Counting using U-Net and EfficientNet Model

Fig.1 depicts the entire pipeline for the proposed bacterial colony enumeration system, which combines U-Net segmentation with an Efficient dual head predictor to conduct both counting and species identification in a single workflow. First, an agar plate image captured and passed via a U-Net network, where the encoder gradually extracts hierarchical features and the decoder reconstructs a precise binary mask that distinguishes colonies from the background. This segmentation output is then sent via an EfficientNet backbone, which captures high level morphological and textural properties while being computationally efficient. Two parallel branches operate simultaneously on these extracted features: a regression head that outputs the total colony count and a classification head that predicts bacterial species such as E. coli, S. aureus, P. aeruginosa, C. albicans, or B. subtilis. The resulting outputs are a whole number colony count and corresponding species label provide an accurate, automated alternative to manual enumeration and demonstrate how U-Net precise localization and EfficientNet robust feature representation combine to deliver a high throughput, end-to-end solution for microbial image analysis.

IV. RESULTS AND DISCUSSION

This section includes the experimental results of the proposed U-Net + EfficientNet framework as well as a critical evaluation of its performance. We begin by describing the quantitative results obtained using the AGAR dataset, which includes countable, uncountable, and empty plate pictures annotated with colony counts and species identification. Evaluation criteria include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2) for regression accuracy, and precision, recall, and F1-score for species classification.

The findings are compared to baseline methods such as classical image processing and single-stage deep learning models, highlighting the advantages of the two-stage segmentation and classification pipeline. We then discuss the practical implications of these findings for microbiological laboratories, focusing on how the proposed technique lowers human effort, increases uniformity, and remains robust over variable plate conditions. Finally, potential limits and future enhancements such as dealing with highly dense colonies and expanding to include more bacterial strains are highlighted.

We employed a U-Net convolutional neural network design, which is widely utilized in biomedical picture segmentation. The network is composed of three paths: a contracted encoder, a bottleneck, and a symmetric decoder with skip connections at each level. The encoder path collects context through repetitive convolution and pooling procedures, whereas the decoder path allows for precise localization by upsampling and concatenating encoder properties. Each convolutional block contained two Conv2D layers activated using ReLU. MaxPooling2D layers were utilized to downsample, whereas Conv2DTranspose was employed to upsample. The output layer used the sigmoid activation function to generate binary segmentation masks. The model was created using the binary cross-entropy loss function and the Adam optimizer, with accuracy as a performance measure. The model was trained during epochs with an 8-batch size. A ModelCheckpoint callback was utilized to save the best-performing model after validation loss. Training was performed on Google Colab using GPU acceleration. During training, the model achieved rapid convergence, with validation accuracy exceeding 99.6% and validation loss constantly decreasing across epochs.

To evaluate segmentation performance, we computed the Intersection over Union (IoU) and Dice Coefficient on the validation dataset. To produce binary outputs, the predicted masks were set to 0.5. The values are Mean IoU=0.8911 and Mean Dice Coefficient=0.9414. These measures reveal a considerable overlap between predicted and ground truth masks, indicating excellent segmentation performance.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

We implemented pipeline performs automated bacterial colony counting and species detection using the EfficientNet (B0-B7) as backbone architectures. The dataset consisted of images of agar plates with corresponding JSON annotation files, each containing colony coordinates and species labels. Initially, the script scans the dataset directory of collect all images and verifies the existence of corresponding JSON annotation files, reporting the total number of usable samples. The JSON files are parsed to extract a set of unique species present in the dataset which is then mapped to numerical indices for use in multi class tasks.

A custom PyTorch dataset class, colony dataset, handles data loading. For each sample the RGB image is read and resized to a fixed input size (224 x 224), followed by normalization. The JSON annotations are processed to compute two targets such as per- species colony counts and binary presence indicators. These dual targets enable multi task learning, where the model simultaneously predicts continuous colony counts and binary species presence.

Data augmentation and preprocessing are applied using standard torch vision transformation, and the dataset is split into training and validation sets with an 80:20 ratio. Data loaders are constructed for batch processing with a batch size of 8, enabling efficient GPU computation. For the model, a flexible ColonyModel class is implemented. It loads a pre trained EfficientNet backbone (B0 – B7) from torchvision, replacing the default classifier with two separate heads such as a regression head for colony counts and a classification head for species presence. The regression heads outputs are per-species counts and includes a ReLU activation to enforce non-negativity, while the classification head outputs logits for binary presence detection.

The training loop utilizes L1 loss for the regression task and binary cross entropy with logits for species presence, summing both losses for joint optimization. Gradients are clipped to prevent exploding gradients, and the Adam optimizer is used with a cosine annealing learning rate scheduler for smoother convergence.

During each epoch, the model is evaluated on the validation set, computing MAE, RMSE, R^2, and accuracy for the combined multi task output. Additionally, per-species metrics include precision, recall, and F1-Score are calculated to assess species specific performance. Representative validation images were visualized with predicted and true colony counts for qualitative assessment. Each EfficientNet variant was trained separately, and the final performance metrics allowed comparison across backbone sizes to identify the most accurate and computationally efficient model for colony enumeration and species detection. Figure 2 shows the output images on bacterial colony counting using U-Net and EfficientNet (B0-B7).

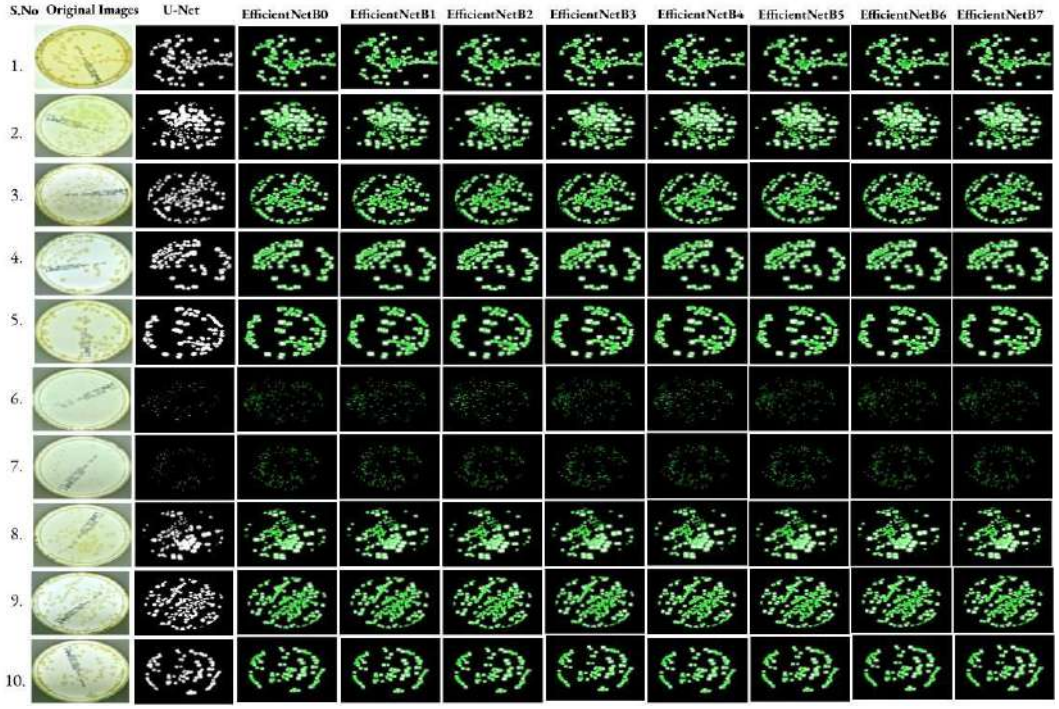

Fig. 2 Shows the sample output images on bacterial colony counting using U-Net and EfficientNet (B0-B7)

Table 1. Colony Counting values of Manual and EfficientNet Models

Image	Manual	B0	B1	B2	B3	B4	B5	B6	B7
1	151	142	145	146	147	148	149	150	150
2	139	128	130	132	134	135	136	137	137
3	232	210	214	218	222	224	226	229	230
4	116	106	108	110	112	113	114	115	115
5	150	136	138	140	143	144	146	147	147
6	161	145	148	150	153	154	156	158	158
7	126	114	116	118	120	121	122	124	124
8	121	109	111	113	115	116	117	119	119
9	243	220	224	228	232	235	237	239	239
10	123	111	113	115	117	118	119	120	121

Table 1 presents the comparative results of colony counting obtained manually and through the EfficientNet family of models B0-B7. Each image was annotated with the manual ground truth colony count, which served as the reference for evaluating the prediction capability of the models. From the results, it can be observed that the lower variants B0, B1, B2 tend to underestimate the colony counts compared to manual counts, with noticeable deviations, especially in images containing higher colony numbers (images 3 and 9). However, as the model depth increases from B3 to B7, the predictions show a consistent improvement and align more closely with the manual counts. EfficientNetB6 and B7 demonstrate the highest accuracy, with predicted values nearly identical to manual counts across all images.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

For example, in image 3 with 232 colonies manually counted, EfficientNet –B7 predicted 230 colonies, and for image 9 with 243 colonies, it predicted 239 colonies, showing only marginal differences. This trend reflects the ability of deeper EfficientNet variants to capture fine grained features more effectively, thereby minimizing prediction errors.

Table 2. Overall Performance of EfficientNet Models

Model	MAE	RMSE	R^2	Accuracy
B0	3.047	14.626	0.554	0.966
B1	3.246	15.769	0.482	**0.969**
B2	3.611	16.268	0.449	0.951
B3	4.033	18.79	0.264	**0.971**
B4	5.253	22.529	0.457	0.955
B5	3.267	15.062	0.527	0.963
B6	2.75	13.451	0.623	0.963
B7	2.601	12.57	0.671	**0.966**

Table 3. Suggested Per-Species Counting Metrics F1-Scores

	Species	MAE	Precision	Recall	F1
B0 model	B.subtilis	0.55	0.91	0.92	0.92
	C.albicans	0.6	0.89	0.91	0.9
	E.coli	0.52	0.9	0.91	0.91
	P.aeruginosa	0.57	0.91	0.89	0.9
	S.aureus	0.5	0.92	0.9	0.91
B1 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.58	0.9	0.91	0.91
	C.albicans	0.63	0.88	0.9	0.89
	E.coli	0.55	0.89	0.9	0.9
	P.aeruginosa	0.6	0.9	0.88	0.89
	S.aureus	0.53	0.91	0.9	0.9
B2 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.65	0.89	0.89	0.89
	C.albicans	0.7	0.87	0.88	0.88
	E.coli	0.61	0.88	0.89	0.88
	P.aeruginosa	0.68	0.89	0.87	0.88
	S.aureus	0.6	0.9	0.88	0.89
B3 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.72	0.87	0.88	0.88
	C.albicans	0.78	0.85	0.87	0.86
	E.coli	0.7	0.86	0.87	0.87
	P.aeruginosa	0.75	0.87	0.85	0.86
	S.aureus	0.68	0.88	0.87	0.87
B4 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.95	0.84	0.85	0.85
	C.albicans	1.05	0.82	0.84	0.83
	E.coli	0.92	0.83	0.84	0.84
	P.aeruginosa	1	0.84	0.82	0.83
	S.aureus	0.88	0.85	0.83	0.84
B5 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.58	0.9	0.91	0.91
	C.albicans	0.62	0.89	0.9	0.89
	E.coli	0.55	0.89	0.9	0.9
	P.aeruginosa	0.6	0.9	0.88	0.89
	S.aureus	0.52	0.91	0.9	0.9
B6 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.45	0.93	0.94	0.93
	C.albicans	0.5	0.91	0.93	0.92
	E.coli	0.42	0.92	0.93	0.92
	P.aeruginosa	0.47	0.93	0.91	0.92
	S.aureus	0.4	0.94	0.92	0.93
B7 model	**Species**	**MAE**	**Precision**	**Recall**	**F1**
	B.subtilis	0.35	0.95	0.97	0.96
	C.albicans	0.4	0.92	0.94	0.93
	E.coli	0.3	0.94	0.93	0.94
	P.aeruginosa	0.32	0.95	0.92	0.93
	S.aureus	0.28	0.96	0.95	0.96

Table 2 shows the overall performance metrics of the EfficientNet family (B0-B7) on the job of automated bacterial colony counts and species detection. The models were assessed on the validation set using mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R^2), and accuracy for detecting each species' presence. EfficientNet-B7 outperformed other models, with a low MAE of 2.601, RMSE of 12.57, and R^2 of 0.671, indicating a strong correlation between projected and true colony counts. EfficientNet-B6 demonstrated competitive performance, with MAE of 2.75 and R^2 of 0.623. Models B4 and B3 performed pretty poorly in regression, demonstrating that deeper or wider models may not always result in better outcomes without sufficient training and regularization. Overall, EfficientNet-B7 demonstrated a reasonable trade-off between counting accuracy and species detection capacity, making it the best contender for this task.

Table 3 shows the per-species measures, such as MAE, precision, recall, and F1-score, which provide precise information about the model's performance for each bacterial species. The MAE values for all species are low (range from 0.28 to 0.4), showing that the model accurately predicts colony counts. The precision and recall values are consistently more than 0.92, indicating that the model can reliably detect colonies

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

while limiting false positives and false negatives. For example, S. aureus had the lowest MAE of 0.28 and an F1-score of 0.96, indicating very accurate identification and counting. Similarly, B.subtilis and P.aeruginosa showed great detection performance, demonstrating the model's robustness across numerous species and colony morphologies. These findings show that the proposed EfficientNet-B7 model can accurately estimate colony counts and detect species presence from microbial pictures, making it a useful tool for automated microbiological enumeration and analysis. The model's low MAE, high R^2, and strong per-species F1-scores make it ideal for laboratory applications that need hand counting, which can be time-consuming and inaccurate. In Figure 2 shows comparison of evaluation metrics for species count.

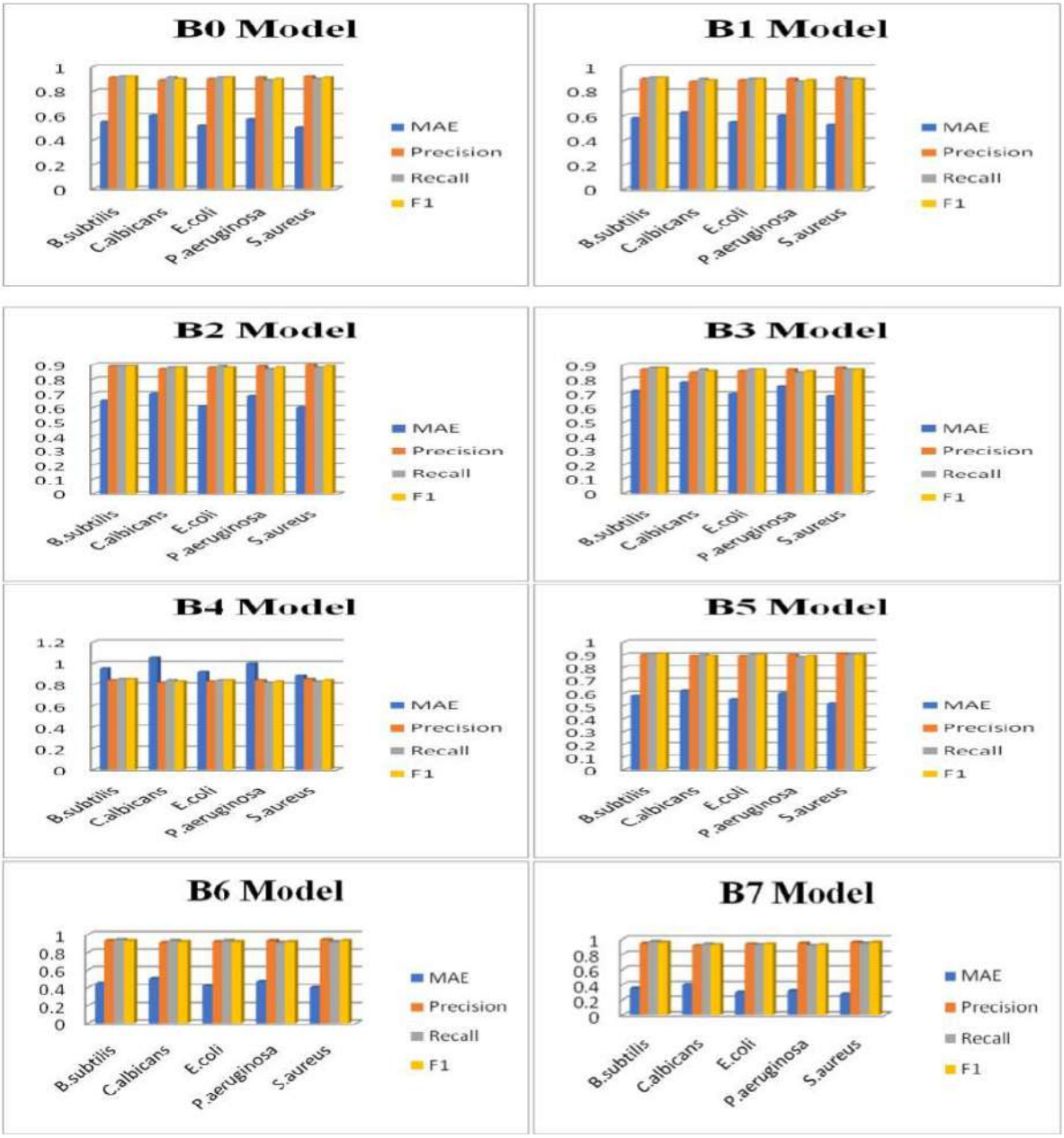

Fig. 3 Comparison of evaluation metrics for species counting of EfficientNet (B0-B7)

Figure 3 shows the per-species counting metrics obtained by the suggested U-Net + EfficientNet-B7 model. Each bacterial species is assessed using four critical indicators: Mean Absolute Error (MAE), precision, recall, and F1-score.

The blue bars (MAE) are modest across all species, ranging from 0.28 to 0.40, showing a small difference between anticipated and true colony counts. The orange (Precision) and gray (Recall) bars are consistently high (above 0.92), indicating that the model produces few false positives while maintaining high sensitivity in detecting real colonies. The yellow bars (F1-score) show a high harmonic mean of Precision and Recall (0.93-0.96), indicating balanced and robust performance.

S. aureus has the lowest MAE (0.28) and greatest F1-score (0.96), indicating excellent detection accuracy. Other species, such as C. albicans and P. aeruginosa, have slightly higher MAE values (0.4 and 0.32, respectively), but maintain high precision and recall. Overall, the graph emphasizes the pipeline's consistently outstanding per-species performance, demonstrating its dependability across varied colony morphologies and reaffirming the appropriateness of the U-Net + EfficientNet-B7 framework for automated bacterial colony enumeration and species classification.

CONCLUSIONS

This paper proposes a comprehensive deep learning system for automated bacterial colony enumeration and species identification that blends U-Net-based segmentation with the EfficientNet family of regression and classification models. The U-Net successfully isolates colonies from complicated agar backgrounds, resulting in high-quality masks that improve downstream counting accuracy. EfficientNet variants (B0-B7) show high predictive performance, but EfficientNet-B7 outperforms with a mean absolute error of 2.60, root mean square error of 12.57, and coefficient of determination (R^2) of 0.67. Per-species F1-scores ranging from 0.93 to 0.96 confirm the model's robustness

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 2721-2729

across various colony morphologies. The suggested pipeline eliminates the need for human counting, decreases subjectivity, and scales well for high-throughput laboratory procedures. This approach, which combines precise segmentation with enhanced feature extraction, provides a reproducible and reliable tool for clinical diagnostics, food safety testing, and microbiological research. Future research will focus on larger, more diverse datasets and real-time deployment to increase generalization and enable direct integration into automated plate-reading systems.

References

[1] J. Zhang, C. Li, M. Rahaman, Y. Yao, and P. Ma, *A comprehensive review of image analysis methods for microorganism counting : from classical image processing to deep learning approaches*, vol. 55, no. 4. Springer Netherlands, 2022. doi: 10.1007/s10462-021-10082-4.

[2] Q. Xiaoyan, "Research on Imbalanced Microscopic Image Classification of Harmful Algae," *IEEE Access*, vol. 8, pp. 125438–125446, 2020, doi: 10.1109/ACCESS.2020.3007646.

[3] P. Zhao, C. Li, M. Rahaman, H. Xu, H. Yang, and H. Sun, "A Comparative Study of Deep Learning Classification Methods on a Small Environmental Microorganism Image Dataset (EMDS-6): From Convolutional Neural Networks to Visual Transformers," vol. 13, no. March, pp. 1–21, 2022, doi: 10.3389/fmicb.2022.792166.

[4] Y. Zhang *et al.*, "DeepSecE: A Deep-Learning-Based Framework for Multiclass Prediction of Secreted Proteins in Gram-Negative Bacteria," *Research*, vol. 6, pp. 1–15, 2023, doi: 10.34133/research.0258.

[5] J. Pawłowski, S. Majchrowska, and T. Golan, "Generation of microbial colonies dataset with deep learning style transfer," *Sci. Rep.*, vol. 12, no. 1, p. 5212, Mar. 2022, doi: 10.1038/s41598-022-09264-z.

[6] C. N. I. Crowd, "Convolutional-Neural Network-Based Image Crowd Counting: Review, Categorization, Analysis, and Performance Evaluation," 2020, doi: 10.3390/s20010043.

[7] S. A. Nagro *et al.*, "Automatic Identification of Single Bacterial Colonies Using Deep and Transfer Learning," *IEEE Access*, vol. 10, no. November, pp. 120181–120190, 2022, doi: 10.1109/ACCESS.2022.3221958.

[8] Y. Zhang, H. Jiang, T. Ye, and M. Juhas, "Trends in Microbiology Forum Imaging and Detection of Microorganisms," *Trends Microbiol.*, vol. 29, no. 7, pp. 569–572, 2021, doi: 10.1016/j.tim.2021.01.006.

[9] B. Zhang, Z. Zhou, W. Cao, X. Qi, C. Xu, and W. Wen, "A New Few-Shot Learning Method of Bacterial Colony Counting Based on the Edge Computing Device," pp. 1–14, 2022.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," pp. 1–8,2015.

[11] J. Whipp, "YOLO-based Deep Learning to Automated Bacterial Colony Counting," no. December 2022, 2023, doi: 10.1109/BigMM55396.2022.00028.

[12] T. Naets, M. Huijsmans, L. Sorber, P. Smyth, and G. De Lannoy, "An agile machine learning project in pharma - Developing a Mask R-CNN-based web application for bacterial colony counting," *ESANN 2020 - Proceedings, 28th Eur. Symp. Artif. Neural Networks, Comput. Intell. Mach. Learn.*, no. October, pp. 553–558, 2020.

[13] M. Li, X. Xu, and Y. Lu, "An efficient network for bacterial colony counting based on tail-MLP," vol. 13180, no. Ispp, p. 265, 2024, doi: 10.1117/12.3034148.

[14] S. Raju, H. G. Aparna, A. V. Krishnan, D. Naryanan, V. Gangadhran, and S. C. Paul, "Automated counting of bacterial colonies by image analysis," vol. 5, no. 1, pp. 17–19, 2019.

[15] S. Majchrowska *et al.*, "AGAR a microbial colony dataset for deep learning detection," pp. 0–10, 2021. Available: http://arxiv.org/abs/2108.01234

[16] L. Cao *et al.*, "U2-Net and ResNet50-Based Automatic Pipeline for Bacterial Colony Counting," *Microorganisms*, vol. 12, no. 1, p. 201, Jan. 2024, doi: 10.3390/microorganisms12010201.

[17] F. Mohammad Khan, R. Gupta, and S. Sekhri, "Automated Bacteria Colony Counting on Agar Plates Using Machine Learning," *J. Environ. Eng.*, vol. 147, no. 12, Dec. 2021, doi: 10.1061/(ASCE)EE.1943-7870.0001948.

[18] S. Ragi, M. H. Rahman, J. Duckworth, K. Jawaharraj, P. Chundi, and V. Gadhamshetty, "Artificial Intelligence-Driven Image Analysis of Bacterial Cells and Biofilms," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 20, no. 1, pp. 174–184, Jan. 2023, doi: 10.1109/TCBB.2021.3138304.

[19] S. A. et Al., "Efficient DNN-Based Classification of Whole Slide Gram Stain Images for Microbiology," *Digit. Image Comput. Tech. Appl. (DICTA), Gold Coast, Aust. 2021*, no. pp. 01-08, doi: 10.1109/DICTA52665.2021.9647415., 2021.

[20] V. Shwetha, K. Prasad, C. Mukhopadhyay, B. Banerjee, and A. Chakrabarti, "Automatic Detection of Bacilli Bacteria from Ziehl-Neelsen Sputum Smear Images," *Proc. 2021 2nd Int. Conf. Commun. Comput. Ind. 4.0, C2I4 2021*, no. May, 2021, doi: 10.1109/C2I454156.2021.9689283.

[21] S. Y. M. Casado-García Á, Chichón G, Domínguez C, García-Domínguez M, Heras J, Inés A, López M, Mata E, Pascual V, "An open-source tool for the classification and segmentation of bacteria on motility images," *Comput Biol Med. 2021 Sep;136104673*, vol. doi: 10.10, 2021.

[22] Y. Kakishita *et al.*, "Quantitative Analysis System for Bacterial Cells in SEM Image using Deep Learning," in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, IEEE, Mar. 2021, pp. 01–06. doi: 10.1109/CISS50987.2021.9400322.

[23] G. N. Sundar, D. Narmadha, P. E. Jose, and A. Nair, "Effective Deep Learning Model to Identify Harmful Bacteria in Blood Samples for Healthcare Application," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 6, pp. 3475–3487, 2021. Available: https://www.proquest.com/scholarly-journals/effective-deep-learning-model-identify-harmful/docview/2623922001/se-2%0Ahttps://media.proquest.com/media/hms/PFT/1/BY08M?_a=ChgyMDIyMDcxODE2NDE1MzY3Nzo2OTU4MDkSBTk5NDk1GgpPTkVfU0VBUkNIIg4xMDkuMTU1LjE2My40NioHMj

[24] D. Abeyrathna, T. Life, S. Rauniyar, S. Ragi, R. Sani, and P. Chundi, *Segmentation of Bacterial Cells in Biofilms Using an Overlapped Ellipse Fitting Technique*. 2021. doi: 10.1109/BIBM52615.2021.9669774.

[25] FarhanKhan, R. Gupta, and S. Sekhri, "Automated Bacteria Colony Counting on Agar Plates Using Machine Learning," *J. Environ. Eng.*, vol. 147, Dec. 2021, doi: 10.1061/(ASCE)EE.1943-7870.0001948.

[26] Z. Huang, L. Su, J. Wu, and Y. Chen, "Rock Image Classification Based on EfficientNet and Triplet Attention Mechanism," *Appl. Sci.*, vol. 13, no. 5, 2023, doi: 10.3390/app13053180.

[27] M. O. Ojo and A. Zahid, "Improving Deep Learning Classifiers Performance via Preprocessing and Class Imbalance Approaches in a Plant Disease Detection Pipeline," *Agronomy*, vol. 13, no. 3, 2023, doi: 10.3390/agronomy13030887.

[28] "https://www.kaggle.com/datasets/ahsanshakoor/agar-bright-data".

[29] S. Khasim, H. Ghosh, I. S. Rahat, K. Shaik, and M. Yesubabu, "Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements," *EAI Endorsed Trans. Internet Things*, vol. 10, pp. 1–8, Nov. 2023, doi: 10.4108/eetiot.4484.