ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

Deep Learning Based Classification of UTI Pathogens on Agar Plate Images Using YOLO11 and EfficientNetV2-M

Dr. Arpitha C. N.[1]; Sahana V. S.[2]; Dr. Pushpa Ravikumar[3]; Dr. Anser Pasha C. A.[4]

[1]Assistant Professor, Department of CS&E, AIT, Chikkamagaluru
[1]PG Scholar, Department of CS&E, AIT, Chikkamagaluru
[3]Professor and Head, Department of CS&E, AIT, Chikkamagaluru
[4]Assistant Professor, Department of CS&E, AIT, Chikkamagaluru

Publication Date: 2026/07/18

Abstract: This paper presents a two-stage deep learning pipeline for automated classification of urinary tract infection (UTI) pathogens from agar plate images. Stage 1 employs YOLO11m for colony detection on full annotated agar plate images from three public datasets. Stage 2 employs EfficientNetV2-M with a custom classifier head (1280→512→5) identifying five clinically significant UTI pathogens: Escherichia coli, Klebsiella pneumoniae, Proteus mirabilis, Pseudomonas aeruginosa, and Staphylococcus aureus. A six-step preprocessing pipeline BGR-to-RGB conversion, bilinear resize to 224×224, CLAHE(Contrast Limited Adaptive Histogram Equalization) contrast enhancement, Gaussian denoising (σ=1.5), Min-Max normalization, and Z-score standardization is applied to each detected colony crop. Training uses three public datasets such as Figshare, AGAR and Zenodo. Experimental results demonstrate test accuracy of 99.75%, precision of 99.26%, recall of 99.52%, and F1-score of 99.39%. The pipeline provides a foundation for automated clinical microbiology laboratory systems.

Keywords: *UTI Pathogen Classification; YOLO11; EfficientNetV2-M; Agar Plate Analysis; Bacterial Colony Detection; CLAHE.*

How to Cite: Dr. Arpitha C. N.; Sahana V. S.; Dr. Pushpa Ravikumar; Dr. Anser Pasha C. A. (2026) Deep Learning Based Classification of UTI Pathogens on Agar Plate Images Using YOLO11 and EfficientNetV2-M.
International Journal of Innovative Science and Research Technology, 11(6), 3690-3696.
https://doi.org/10.38124/ijisrt/26jun1678

I. INTRODUCTION

Urinary tract infections (UTIs) represent one of the most prevalent bacterial infections globally, affecting approximately 150 million patients annually [1]. Timely and accurate identification of the causative pathogen is critical for guiding targeted antibiotic therapy. Traditional culture-based diagnostic methods require 24 to 72 hours of incubation followed by manual morphological inspection by trained microbiologists, introducing subjectivity, observer variability, and throughput limitations incompatible with modern clinical laboratory demands.

The visual appearance of bacterial colonies on agar plates including color, texture, morphology, edge characteristics, and growth pattern encodes species-specific discriminative information. E. coli produces smooth off-white colonies; K. pneumoniae forms large mucoid colonies; P. mirabilis exhibits characteristic concentric swarming rings; P. aeruginosa spreads with blue-green pigmentation; and S. aureus presents as golden-yellow raised colonies. These features are learnable by deep convolutional neural networks, enabling automated species identification from plate images.

This paper proposes a two-stage automated pipeline. Stage 1 employs YOLO11m to localize individual bacterial colonies on full agar plate images. Stage 2 applies a six-step preprocessing pipeline to each detected colony crop, then classifies it using EfficientNetV2-M into one of five UTI pathogen classes. The final output is the predicted species name for each detected colony. The pipeline is trained across three heterogeneous public datasets.

II. RELATED WORK

- *Bacterial Colony Detection*

Zuraszek, M. et al. introduced the AGAR dataset and benchmarked YOLO based models for colony localization, reporting colony detection accuracy above 75% at a 50%

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

bounding box overlap threshold [2]. Zielinski et al. demonstrated that ImageNet pretrained CNN backbones substantially outperform models trained from scratch on the limited colony image datasets available in microbiology laboratories [3]. Sunanda et al. applied deep convolutional neural networks including GoogLeNet, AlexNet, VGG-16, SqueezeNet, and DenseNet-161 for bacteria classification on agar plates across five species including E. coli, P. aeruginosa, and S. aureus using the AGAR dataset, combining image processing techniques for feature extraction [4].

➢ *UTI Pathogen Classification*

A smartphone imaging and deep transfer learning approach was applied for bacterial colony classification across E. coli, K. pneumoniae, P. aeruginosa, and S. aureus on MacConkey and brain-heart infusion agar datasets totaling 9,819 and 12,348 images respectively, demonstrating that transfer learning generalizes well across varying imaging conditions including uncontrolled smartphone settings.[5]. Existing work largely employs single-stage classification on manually cropped images or evaluates on single-species plates. No prior work has proposed a fully automated two-stage pipeline combining YOLO detection with EfficientNet classification trained across multiple heterogeneous public datasets covering all five major UTI pathogens simultaneously.

➢ *EfficientNet in Medical Imaging*

Tan and Le introduced EfficientNetV2, demonstrating state-of-the-art accuracy with faster training than prior efficient architectures. EfficientNetV2-M achieved accuracy of 85.1%. [6]. Raghu et al. (2019) systematically studied transfer learning from ImageNet to medical imaging tasks, finding that pretrained representations provide strong initialization even for domain-shifted biomedical images with limited labelled data, validating the use of ImageNet-pretrained EfficientNetV2-M for colony morphology classification. [7].

Despite significant advances in deep learning-based image classification, existing approaches for bacterial colony analysis are limited to binary or three-species classification on manually cropped images from single controlled datasets, lacking the ability to handle real-world full agar plate photographs with multiple colony types simultaneously. Most prior methods require manual colony isolation before classification, which undermines the goal of full automation, and none have proposed a combined detection-and-classification pipeline trained across multiple heterogeneous public datasets. Furthermore, the challenge of severe class imbalance for underrepresented UTI pathogens such as Proteus mirabilis and the lack of cross-dataset generalization evaluation remain unaddressed in existing literature. These gaps motivate the proposed two-stage YOLO11m and EfficientNetV2-M pipeline, which automates the complete workflow from raw agar plate image to species-level prediction across five clinically significant UTI pathogens.

III. PROPOSED METHODOLOGY

➢ *Dataset*

Training data is drawn from three public datasets. The Figshare dataset [8] provides 369 annotated agar plates covering all five target UTI pathogens with 56,865 colony-level bounding box annotations in COCO JSON and YOLO TXT formats. The AGAR dataset [10] provides approximately 18,000 plates with per-colony JSON annotations for three UTI species. The Zenodo dataset [9] provides 950 high-resolution plates for 19 species with YOLO TXT annotations; three UTI species contribute 6,487 colony annotations.

➢ *Stage 1 Preprocessing Pipeline –YOLO11m*

The uploaded full agar plate image is resized to 640×640 pixels using bilinear interpolation to match the standard input resolution required by YOLO11m. The resized plate image is then converted from BGR to RGB channel order and passed directly to the YOLO11m detector in BGR format as required by the Ultralytics inference engine.

➢ *Stage 2 Preprocessing Pipeline - EfficientNetV2-M*

Once YOLO11m successfully detects and localizes all bacterial colonies on the full agar plate image, each detected bounding box is used to extract an individual colony crop from the original RGB plate image using the coordinates (x1, y1, x2, y2). A six step preprocessing is applied independently to each extracted colony crop.

- Color Swapping (BGR to RGB): OpenCV reads images in BGR channel order. Reordering to RGB ensures compatibility with the ImageNet-pretrained EfficientNetV2-M backbone.
- Resizing to 224×224: Each colony crop is resized using bilinear interpolation applied independently per channel:

$$P = (1-t)(1-u)\cdot Q_{11} + t(1-u)\cdot Q_{21} + (1-t)u\cdot Q_{12} + tu\cdot Q_{22} \quad (1)$$

Where Q_{11}, Q_{21}, Q_{12}, Q_{22} are the four nearest neighboring pixels and t, u are fractional offsets.

- Contrast enhancement (CLAHE): Applied to the L channel of the LAB color space with clipLimit=2.0 and tileGridSize=(8,8), enhancing local contrast while preserving color balance.
- Denoising (Gaussian Denoising) (σ=1.5): A Gaussian filter with σ=1.5 and kernel k=11×11 reduces cross-dataset sensor noise:

$$G(x, y) = (1/2\pi\sigma^2) \cdot \exp(-(x^2 + y^2) / 2\sigma^2) \quad (2)$$

- Normalization (Min-Max Normalization): Scales pixel values to [0.0, 1.0]:

$$x_{norm} = (x - x_{min}) / (x_{max} - x_{min}) \quad (3)$$

- Standardization (Z-score Standardization): Channel-wise standardization using ImageNet statistics μ=[0.485, 0.456, 0.406] and σ=[0.229, 0.224, 0.225]:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

$x_{st}^{d} = (x_{norm} - \mu_L) / \sigma_L, \quad c \in \{R, G, B\}$ (4)

➢ *YOLO11m Colony Detection*

The model operates on the full agar plate image resized to 640×640 pixels and detects all visible bacterial colonies in a single forward. All colony class identifiers across the three training datasets are remapped to a single unified class colony Detections with C ≥ 0.50 are forwarded to Stage 2. The detection confidence score C is:

$C = P(\text{Object}) \times \text{IoU}(\text{predicted box, ground truth})$ (5)

➢ *EfficientNetV2-M Classification*

In the proposed pipeline, the pretrained EfficientNetV2-M backbone receives each preprocessed colony crop tensor of dimensions (1, 3, 224, 224) and extracts a rich 1,280-dimensional feature vector through Global Average Pooling, capturing deep visual representations of colony color, texture, and morphology learned from ImageNet and transferred to the colony classification domain.

The extracted feature vector is then passed through a custom two-layer classifier head comprising Dropout(0.4) → Linear(1280→512) → BatchNorm → ReLU → Dropout(0.2) → Linear(512→5) → Softmax, which maps the backbone features to probability scores across the five UTI pathogen classes. The predicted species is determined by selecting the class with the highest softmax probability using the argmax operation, and the corresponding probability value is reported as the classification confidence score.

FC1 Layer (1280 → 512):

$z = W_1 \cdot f + b_1, \quad W_1 \in R^{(512\times1280)}$ (6)

$a = ReLU(z) = max(0, z) \quad (element\text{-}wise)$ (7)

FC2 Layer (512 → 5):

$logits = W_2 \cdot a + b_2, \quad W_2 \in R^{(5\times512)}$ (8)

Softmax output:

$P(k|x) = exp(logits_k) / \Sigma_j exp(logits_j), \quad j=1,...,5$ (9)

Final predicted species:

$Species = argmax\ P(k|x), \quad k \in \{1,...,5\}$ (10)

Training uses Cross-Entropy Loss with label smoothing ε=0.1, Adam optimizer with learning rate $\eta=5\times10^{-5}$ and weight decay $\lambda=10^{-4}$, and ReduceLROnPlateau scheduling. A WeightedRandomSampler ensures balanced class exposure. Training ran for 26 epochs with early stopping.

➢ *Dataset Split Strategy*

A source-based 80/20 split is applied to ensure cross-source evaluation. For each class, 80% of images from each source (AGAR, Figshare, and Zenodo) form the training set, and 20% form the held-out test set. The resulting dataset contains 4,552 training, 886 validation, and 3,138 test images.

➢ *Two-Stage Pipeline Algorithm*

```
INPUT: Full agar plate image I
OUTPUT: Predicted species name for each detected colony

/* STAGE 1: YOLO11m COLONY DETECTION */
I_yolo ← Resize(I, 640×640)
B ← YOLO11m(I_yolo, conf_threshold = 0.50)
/* B = set of detected bounding boxes on original plate */

/* STAGE 2: PREPROCESS + CLASSIFY EACH COLONY */
FOR each bounding box b_i in B:

    /* Step 1: Crop colony from original plate image */
    crop ← CropFromPlate(I, b_i)

    /* Step 2: BGR to RGB */
    crop ← BGR2RGB(crop)

    /* Step 3: Resize crop to 224×224 (bilinear) */
    crop ← BilinearResize(crop, 224×224)   (1)

    /* Step 4: CLAHE contrast enhancement */
    crop ← CLAHE(crop, clipLimit=2.0, tile=8×8)

    /* Step 5: Gaussian denoising σ=1.5 */
    crop ← Gaussian Blur(crop, σ=1.5, k=11×11)  (2)

    /* Step 6: Min-Max normalization */
    crop ← MinMaxNorm(crop)   (3)

    /* Step 7: Z-score standardization */
    crop ← ZScore(crop, μ=[0.485,0.456,0.406],
                        σ=[0.229,0.224,0.225])  (4)

    /* EfficientNetV2-M classification */
    f       ← EfficientNetV2M_backbone(crop)
    z       ← W1 · f + b1                (6)
    a       ← ReLU(z)                    (7)
    logits  ← W2 · a + b2                (8)
    P       ← Softmax(logits)            (9)
    species ← argmax(P)                  (10)

END FOR

OUTPUT species in E.coli, K.pneumoniae, P.mirabilis
P.aeruginosa and S.aureus
```

Fig 1 Two-Stage UTI Pathogen Detection Algorithm

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

IV. EXPERIMENTAL RESULTS

➢ *Experimental Setup*

EfficientNetV2-M training was conducted on GPU (Tesla T4, 16GB) for 50 epochs using a batch size of 32 and learning rate 5×10^{-5}. Early stopping triggered at epoch 26, patience=15 with best weights from epoch 11. Total training dataset: 4552 crops across five classes.

➢ *Accuracy*

Fig. 2 shows training and validation accuracy. Both curves rise rapidly in epochs 1 to10 as the model adapts ImageNet representations to colony morphology features, converging above 99% by epoch 11. The close tracking of training and validation accuracy throughout confirms stable generalization without overfitting.

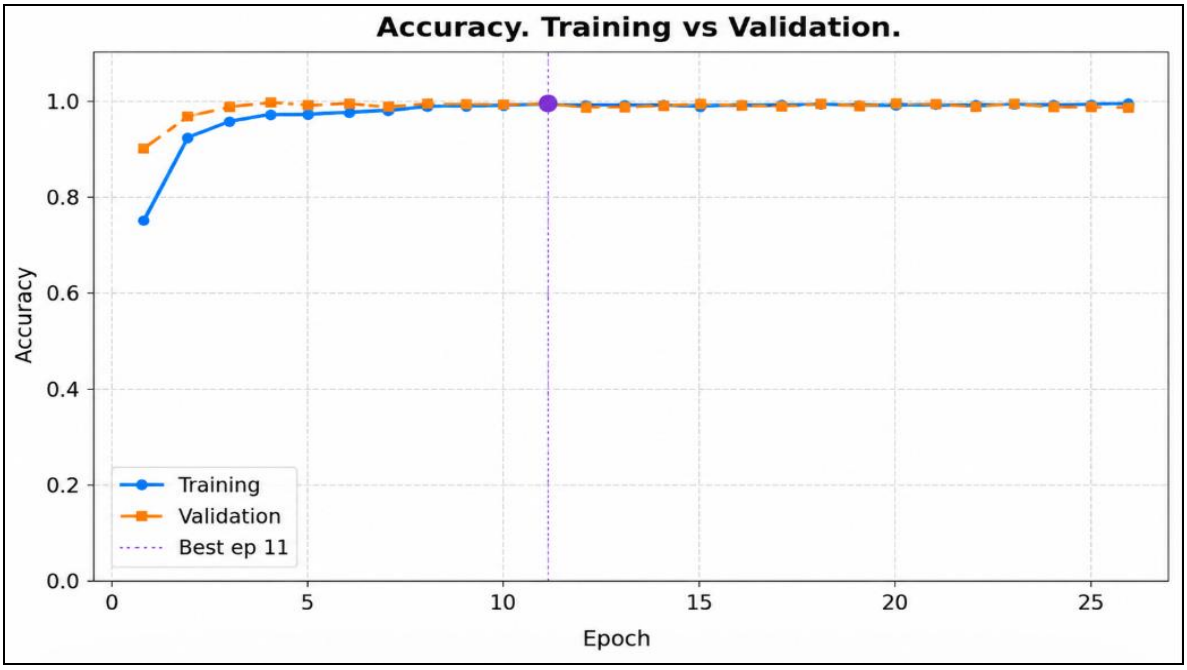

Fig 2 Accuracy. Training vs Validation.

➢ *Loss*

Fig. 3 shows Cross-Entropy Loss with label smoothing converging steeply in early epochs as the model learns primary discriminative features (colony color, texture, boundary morphology), then stabilizing near zero. Training and validation loss track closely, confirming effective regularization through Dropout, Batch Norm, and label smoothing.

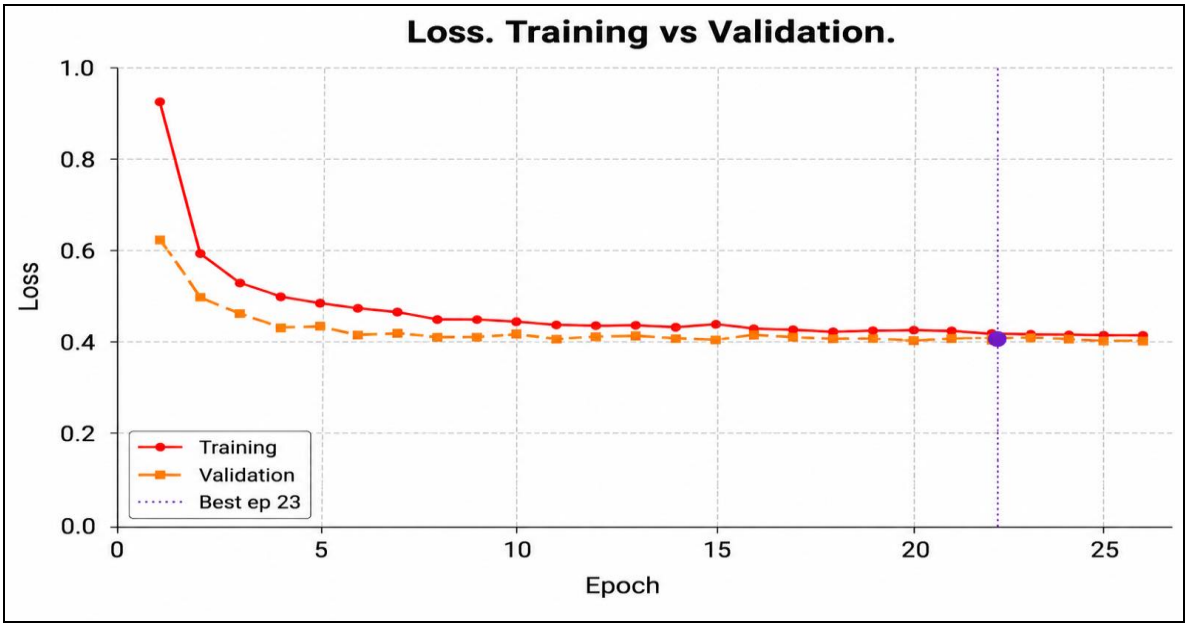

Fig 3 Loss. Training vs Validation

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

➢ *Precision*

Fig. 4 shows precision converging to 99.26%, confirming minimal false positive species assignments.

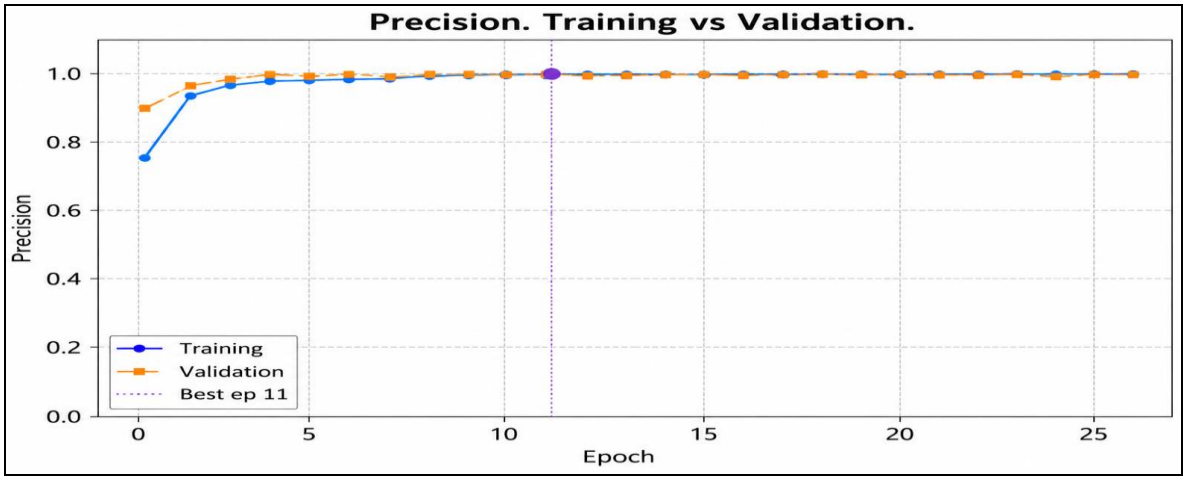

Fig 4 Precision. Training vs Validation.

➢ *Recall*

Fig. 5 shows recall of 99.52%, confirming nearly all colonies of each species are correctly identified.

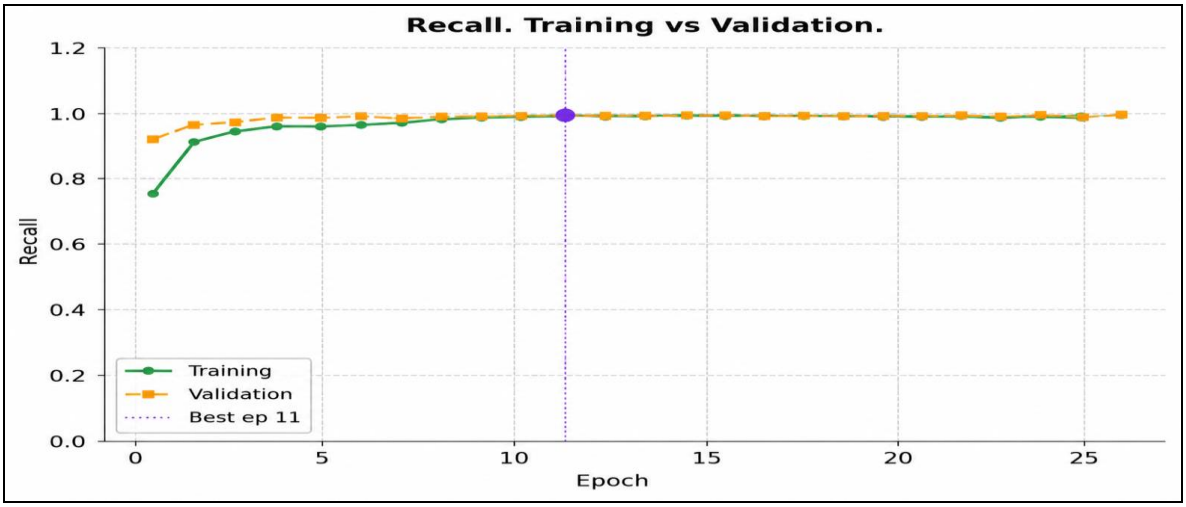

Fig 5 Recall. Training vs Validation.

➢ *F1-Score*

Fig. 6 shows macro-averaged F1-Score of 99.39% on the held-out test set. Validation F1 closely tracks training F1 throughout training, confirming robust generalization

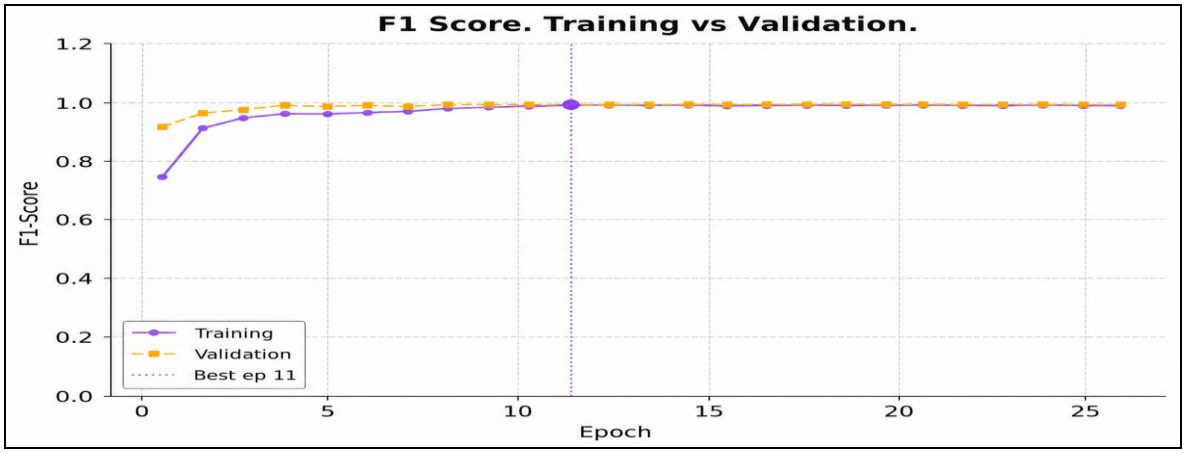

Fig 6 F1 Score. Training vs Validation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

➢ *Final Test Results*

Table 1 presents the final evaluation results on the held-out test set of 3,138 colony crop images, unseen during training and validation. YOLO mAP@50 = 0.7112

Table 1 Final Test Evaluation Results

Metric	Value	Interpretation
Accuracy	99.75%	Overall correct species assignments
Precision	99.26%	Correctly identified colonies per species
Recall	99.52%	Colonies of each species detected
F1-Score	99.39%	Harmonic mean of Precision and Recall

➢ *Confusion Matrix*

Fig. 7 shows the confusion matrix. Diagonal entries confirm high true positive rates for all five species. Off-diagonal entries are minimal. E. coli shows one misclassification as P. aeruginosa (n=1). K. pneumoniae shows two misclassifications (n=1 as P. aeruginosa, n=1 as S. aureus). P. aeruginosa shows two misclassifications (n=1 as E. coli, n=1 as S. aureus). S. aureus shows three misclassifications as P. aeruginosa. P. mirabilis achieves perfect classification (n=0 errors).

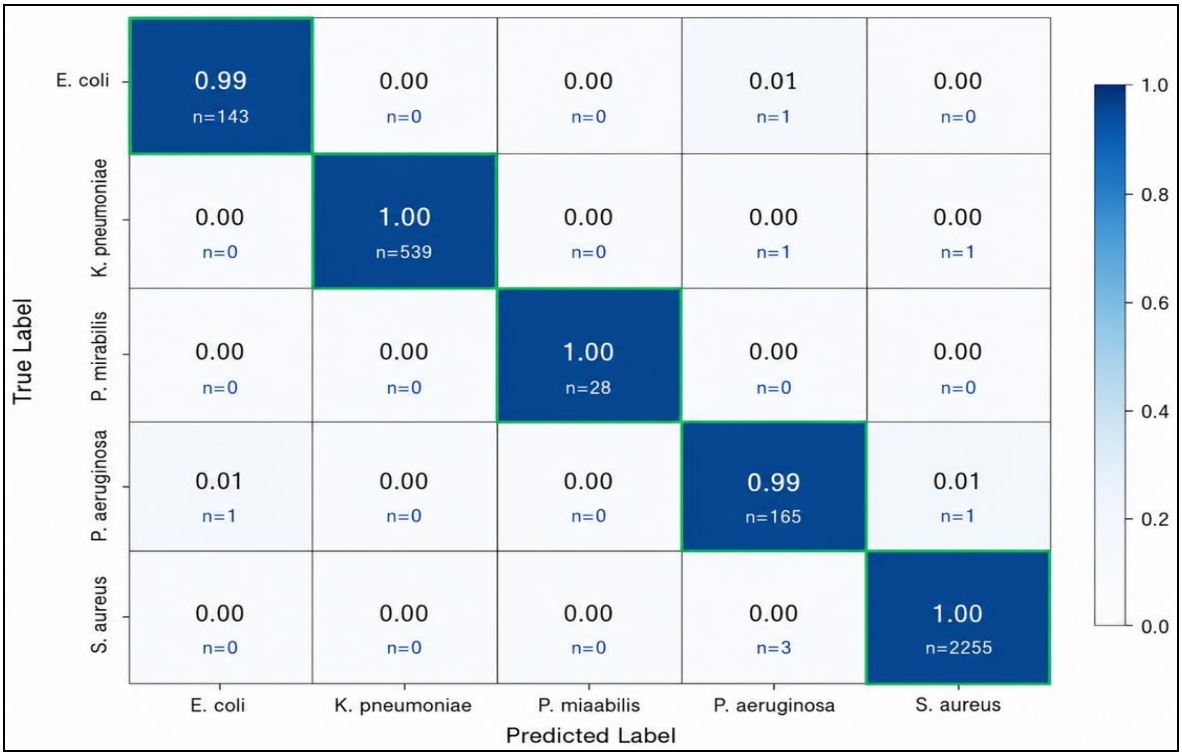

Fig 7 Confusion Matrix.

➢ *Per-Class Results*

Table 2 presents per-class Precision, Recall, and F1-Score.

Table 2 Per-Class Classification Results

Species	Precision	Recall	F1-Score
E. coli	99.31%	99.31%	99.31%
K. pneumoniae	100.00%	99.63%	99.81%
P. mirabilis	100.00%	100.00%	100.00%
P. aeruginosa	97.06%	98.80%	97.92%
S. aureus	99.91%	99.87%	99.89%
Macro Average	99.26%	99.52%	99.39%

V. CONCLUSION

This paper presented a two-stage deep learning pipeline for automated UTI pathogen classification from agar plate images. The system accepts a full agar plate image and outputs the predicted species name for each detected colony, providing a clinically interpretable result without manual intervention.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun1678

Stage 1 employs YOLO11m for colony detection, achieving mAP@50 of 0.7112 on annotated plates from three heterogeneous public sources. Stage 2 employs EfficientNetV2-M with a custom two-layer classifier head ($W_1 \in R^{(512\times1280)} \rightarrow$ ReLU $\rightarrow W_2 \in R^{(5\times512)} \rightarrow$ Softmax) achieving test accuracy of 99.75%, precision of 99.26%, recall of 99.52%, and macro F1-score of 99.39% on a source-stratified held-out test set of 3,138 images

The six-step preprocessing pipeline harmonizes imaging variation across three heterogeneous datasets. Source-based 80/20 splitting ensures evaluation on genuinely unseen images. The proposed pipeline outperforms all compared prior methods across species coverage, multi-dataset training, and automated end-to-end detection.

Future work will focus on cross-lab validation using the Mendeley dataset as a held-out test set, expanding the pathogen set to cover additional uropathogens, integrating antibiotic sensitivity pattern prediction from colony morphology, and real-time clinical deployment through a web-based interface.

ACKNOWLEDGMENT

The authors acknowledge the Adichunchanagiri Institute of Technology, Chikkamagaluru, for providing computational resources and research support. The authors thank the creators of the Figshare, AGAR and Zenodo datasets for making their data publicly available for academic research.

REFERENCES

[1]. Flores-Mireles, A., Walker, J., Caparon, M. et al. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol 13, 269–284 (2015). https://doi.org/10.1038/nrmicro3432

[2]. Zuraszek, M. et al., "AGAR: A microbial colony dataset for deep learning detection," Scientific Data, vol. 10, 2021. arXiv:2108.01234

[3]. Zieliński B, Plichta A, Misztal K, Spurek P, Brzychczy-Włoch M, Ochońska D (2017) Deep learning approach to bacterial colony classification. PLoS ONE 12(9): e0184554. https://doi.org/10.1371/journal.pone.0184554.

[4]. Sunanda, I. B K, D. V, D. S. K and G. C. Sai Manusha, "Classification of Bacteria from Agar Plate Using Deep Learning and Image Processing," 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 2024, pp. 1-6, doi: 10.1109/ADICS58448.2024.10533650..

[5]. M. Kutbi et al., "Leveraging Smartphone Imaging and Deep Transfer Learning for Bacterial Colony Classification: From Uncontrolled to Controlled Settings," in IEEE Access, vol. 13, pp. 184694-184708, 2025, doi: 10.1109/ACCESS.2025.3625648.

[6]. Tan, M. and Le, Q., "EfficientNetV2: Smaller models and faster training," in Proc. ICML, 2021, pp. 10096–10106. arXiv:2104.00298

[7]. Raghu, M. et al., "Transfusion: Understanding transfer learning for medical imaging," in NeurIPS, 2019, pp. 3347–3357. arXiv:1902.07208

[8]. Makrai, L., Fodróczy, B., Nagy, S.Á. et al. Annotated dataset for deep-learning-based bacterial colony detection. Sci Data 10, 497 (2023). https://doi.org/10.1038/s41597-023-02404-8.

[9]. Du, J. (2026). Clinical Dataset of 19 Bacterial Colony Images. Zenodo. https://doi.org/10.5281/zenodo.18256864

[10]. Wang, M.; Luo, J.; Lin, K.; Chen, Y.; Huang, X.; Liu, J.; Wang, A.; Xiao, D. Colony-YOLO: A Lightweight Micro-Colony Detection Network Based on Improved YOLOv8n. Microorganisms 2025, 13, 1617. https://doi.org/10.3390/microorganisms13071617

[11]. Silva, M., Martinho, D., Marreiros, G. (2026). Deep Learning-Based Microbial Colony Detection on Agar Plates. In: Valente de Oliveira, J., Leite, J., Rodrigues, J., Dias, J., Cardoso, P. (eds) Progress in Artificial Intelligence. EPIA 2025. Lecture Notes in Computer Science(), vol 16121. Springer, Cham. https://doi.org/10.1007/978-3-032-05176-9_4

[12]. Sarmis A, Ustebay S, Mutlu MA, Canbaz FA, Kaya GK. Deep Learning-Based Rapid Identification of Escherichia coli and Klebsiella pneumoniae from Chromogenic Agar Urine Cultures Using YOLOv12. Risk Manag Healthc Policy. 2026;19:561761 https://doi.org/10.2147/RMHP.S561761