

Identification of Microbial Strains Using an Instance Segmentation Model

Carter Davis¹, Alexey Lange², Sadie Blaszczyk³, Heidi Heldt³, Nicole Vander Schaaf³, José Manjarrés²

¹Department of Mathematical, Information and Computer Sciences, Point Loma Nazarene University, San Diego, USA

²Department of Physics and Engineering, Point Loma Nazarene University, San Diego, USA

³Department of Biological Sciences, Olivet Nazarene University, Bourbonnais, Illinois, USA

{cdavis0022, alange0022, jmanjarr}@pointloma.edu {navanderschaaf, slblaszczyk, hmheldt}@olivet.edu

Abstract—Accurate identification of microbial strains in urinary tract infections (UTIs) is crucial for effective treatment and prevention of severe complications, such as kidney damage and mortality. Diagnostic methods used in high-resource countries are often not feasible in rural medical institutions causing a delay in diagnosis and treatment. This study addresses this issue by developing a cost-efficient and computationally inexpensive alternative using a customized YOLOv8 (You Only Look Once) deep learning model for microbial identification. UTIs are most effectively treated when the particular microbial strain is identified, making accurate detection essential. We implemented the Fast Segment Anything Model (FastSAM), a powerful segmentation model, to create masks serving for the training of the YOLOv8 model. Our approach was evaluated on a dataset of nine common microbial strains: *Candida albicans*, *Enterococcus faecalis*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus saprophyticus*, and *Streptococcus agalactiae*. The model achieved a notable mean average precision of 0.79, which demonstrates its ability to detect and identify microbial strains. The results are indicative of the model's potential as a reliable and accessible diagnostic aid. This approach can make microbial identification more efficient and widely available, thus improving patient treatment with accurate and timely diagnosis.

Index Terms—YOLOv8, FastSAM, Instance Segmentation, Bacteria Identification.

I. INTRODUCTION

Urinary tract infections (UTIs), if left untreated, can lead to health complications like permanent kidney damage or septic shock [1]. Each year, 150 million people are affected by UTI, predominantly impacting women, children, and elders [2]. Easy and accurate bacterial strain identification is crucial for effectively treating these infections. While well-equipped medical institutions often have the means to utilize high-performance equipment for bacterial identification, many healthcare facilities lack access to such technology. An estimated 70% of medical equipment designed in developed countries is unsuitable for developing regions due to a lack of stable energy sources, trained personnel, or supporting equipment [3]. Consequently, these facilities resort to empirical methods, often delaying accurate diagnosis and treatment.

A solution to the lack of specialized equipment in developing countries is the ability to take photos of petri dishes with infected urine samples and determine the bacterial strain using a mobile device. In recent years, there has been significant progress in the development of mobile-friendly

computer vision models. YOLOv8 is a newer iteration of YOLO (You Only Look Once) [4], a light computer vision model that performs object detection, that improves on its predecessors by including instance segmentation along with image classification. YOLOv8 is fast to train, easy to develop, and suitable for use on computationally-limited devices [5]. We use this model to detect nine bacteria including *Candida albicans*, *Enterococcus faecalis*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Staphylococcus saprophyticus*. We chose these bacteria strains as they are commonly found in human urine samples. Cultures in our study were grown over a 72-hour period. Since different strains of bacteria exhibit varying growth patterns over time, we took pictures at different intervals after the start of incubation.

We utilized the YOLOv8 model in conjunction with FastSAM, the Fast Segment Anything Model. FastSAM is a transformer-based model, which learns from images divided into small sections, derived from the Segment Anything Model (SAM) that generates precise segmentation masks for all instances within an image by learning from images divided into small sections [6]. Segmentation masks are regions returned by the model indicating what it predicts is a particular object in the image [7]. The input data for the YOLOv8 model comprises the masks generated by FastSAM and the corresponding labels for each bacterial strain.

Previous work has demonstrated the potential of machine learning for diagnosing bacterial pathogens in urine samples. Andreini et al. show success using automated systems to capture digital images of petri dishes seeded with infected samples, preprocessing these images to isolate bacterial colonies, and then classifying the infection type using Support Vector Machines (SVMs) [8]. In our study, we implement YOLO rather than SVM, as YOLO is faster due to its single-stage architecture which processes the entire image in one pass, whereas SVM requires separate stages for feature extraction and classification, causing slower performance [9]. A study by Wang et al. uses advanced data collection methods alongside neural networks to successfully make real-time bacteria detections. Through the integration of an incubator, platform scans, and analyzing the plate surface every 30 minutes, they successfully detected bacterial colonies [10]. This study performs bacterial identification on only three strains of bac-

teria, whereas our research aims to identify a greater range of bacteria. Our study aims to perform identification with a simple image rather than the need for costly equipment such as an incubator or platform scanner.

Colony morphology refers to the visual appearance of bacterial colonies on an agar plate and is pertinent knowledge for identifying a bacterial pathogen. [11]. It is critical to observe and characterize colony morphology to effectively identify microorganisms, thus we choose to implement an instance segmentation model which allows for the isolation and analysis of individual colonies within a petri dish for classification purposes. Instance segmentation generates a pixel-by-pixel segmentation mask of the precise shape and area of each instance [12]. This precise segmentation is valuable when creating a model that can accurately characterize colony morphology, allowing for the identification of individual colonies and their particular strain.

The contributions of this paper include, 1) developing a method to detect and segment bacterial colonies within images using FastSAM, allowing for detailed examination and isolation of individual colonies, and 2) implementing a YOLOv8 model for the classification of individual bacterial strains, demonstrating its efficacy in identifying different bacterial colonies and highlights its potential as a diagnostic tool.

The structure of this paper is as follows. Section 2 describes the methodology spanning from data collection to training of our model. In Section 3, the results of the model are presented. Section 4 includes a discussion of the model and results. Section 5 is a conclusion to the study.

II. METHODOLOGY

Figure 1 provides an overview of the three stages of our work. First, we collect images for our dataset. Then, we obtain segmentation masks using FastSAM and create training and validation data. Finally, we train and validate our instance segmentation model.

A. Data Collection

We choose a dataset comprised of images of nine bacterial strains, as shown in Table I. We cultivate each bacteria strain individually in petri dishes. These strains are planted using sterile techniques, including the use of biosafety level 2 cabinets for handling pathogenic organisms.

We selected Tryptic soy agar (TSA) as the media type as it was compatible for each microorganisms growth. We incubate samples at 37°C and remove them from incubation to collect images at four time intervals: 18, 24, 48, and 72 hours. We take 15 images of each dish at each time interval to provide a robust and diverse data set to enhance the model's generalization, or ability to generalize its existing recognized patterns on new data, capabilities [12].

B. Data Preparation

To train the model on adequate and appropriate images, we perform data preparation. During our initial visual analysis, we observe that images taken at the 18-hour mark often

exhibited minimal bacterial growth, with colonies too small for effective detection. Conversely, images we took at the 72-hour mark frequently exhibited overgrowth making individual colony identification difficult (Fig. 2). We manually remove these images during the data preparation process of each strain to ensure the dataset's quality and relevance. Table II shows the number of images we remove for each strain.

We manually crop every image to eliminate background objects that could negatively impact the model's training. Additionally, some images were either too blurry for clear colony detection and thus were removed, further polishing the dataset.

To perform labeling of particular bacteria colonies in each image, there were two methods. One solution would be to manually label colonies in each image, but the process of manually labeling was deemed time-consuming. As an alternative, we automate the labeling process and identification of the colonies using FastSAM. FastSAM segments objects within an image and has the functionality of assigning labels to them. Fig. 3 displays the architecture of the FastSAM model. It consists of a transformer-based model used to segment each image, along with a Feature Pyramid Network (FPN) that is used to assign labels to its detections. An FPN aids in the analysis of images from different perspectives, meaning focused on smaller or larger portions of the image.

We provide bacteria labels to FastSAM to apply to the segmentation masks it created. The labels for the bacteria range from 0 to 8, each representing a different bacterial strain. For instance segmentation, FastSAM uses existing masks as prompts and will select the region with the highest Intersection over Union (IoU) as the output mask. Equation (1) represents this calculation, where A is a predicted area or mask and B is a ground truth area or mask within the image. The area where A and B intersect is divided by the area where either A and B may lie in the picture.

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Fig. 4 includes a visualization of the process of FastSAM overlaying masks onto each image where it detected an object. After FastSAM processes an image, the X and Y coordinates outlining the perimeter of each mask are returned by the model. Subsequently, we transfer these coordinates into a text file, sharing the same name as the image.

Next, we run each file through a function used to filter out any masks that were unlikely to be colonies. We choose a value of 300 data points per mask as the cutoff for objects that seemed too large to be deemed a mask. If no valid masks remain in a file, we remove it from the dataset along with its corresponding image. Once we filter the data, we pair each text file with its image by creating a function to traverse the files, and if the names match we place them in a new folder together.

For the final step in data preparation, we separate the data pairs into a training and validation set. The training set contains 90% of the total data, which is 1,776 images. The

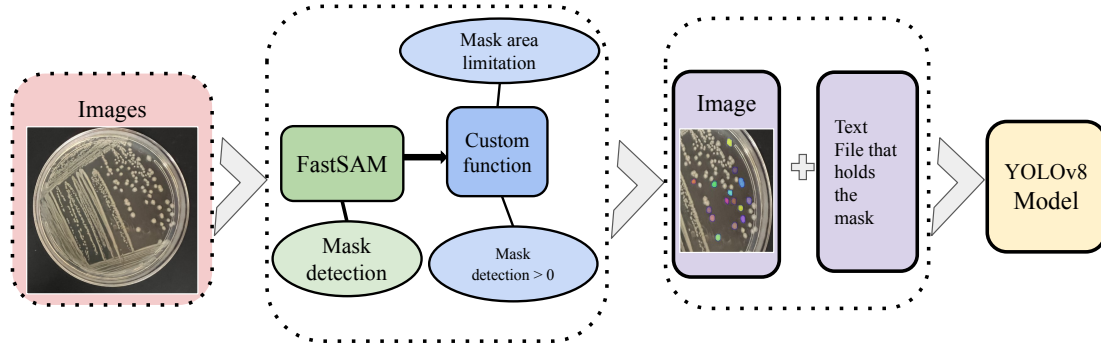


Fig. 1. Illustration of our methodology, including the automatic mask generation and training of the instance segmentation model.

TABLE I
THE NINE STRAINS OF BACTERIA CHARACTERISTICS

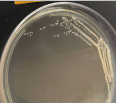
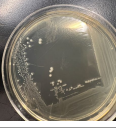
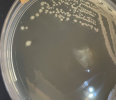
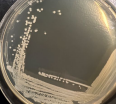
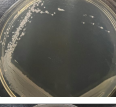



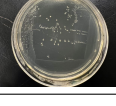
Strain of Bacteria	Image	Characteristics
<i>Candida albicans</i>		<i>Candida albicans</i> colonies have a white, pearl-like appearance. In the fourth quadrant, there are very few colonies, while in the third quadrant, the colonies form a characteristic "pearl chain".
<i>Enterococcus faecalis</i>		<i>Enterococcus faecalis</i> colonies are clear and extremely small. In the fourth quadrant, the tiny colonies grow in rows.
<i>Escherichia Coli</i>		<i>Escherichia coli</i> colonies are milky white and condensed, appearing larger in size with a rhizoid form. The colonies are not prevalent in the middle of the culture.
<i>Klebsiella pneumoniae</i>		<i>Klebsiella pneumoniae</i> colonies are characterized by a milky, white appearance and a glossy, smooth surface.
<i>Pseudomonas aeruginosa</i>		<i>Pseudomonas aeruginosa</i> colonies often appear greasy with a green tint. They form numerous clusters and are less circular, often exhibiting a spindle shape.
<i>Staphylococcus aureus</i>		<i>Staphylococcus aureus</i> colonies are glossy and white, ranging in size from medium to tiny. They are convex and round. The third set of colonies is typically tiny and forms cluster trains, while the fourth set is scattered, medium-sized, and clear.
<i>Staphylococcus saprophyticus</i>		<i>Staphylococcus saprophyticus</i> in the third quadrant forms small, tiny circular colonies. In the fourth quadrant, the colonies remain small and tiny but clump into a zigzag pattern. Occasionally, they spread out into larger white circular colonies.
<i>Staphylococcus epidermidis</i>		<i>Staphylococcus epidermidis</i> colonies are extremely tiny and small, with a white, round, and glossy appearance. There is no growth in the center and little in the fourth quadrant.
<i>Streptococcus agalactiae</i>		<i>Streptococcus agalactiae</i> colonies are clear and white, with extremely tiny circular colonies. In the third quadrant, the colonies are very tiny, while in the fourth quadrant, they form either about five white colonies spread apart or a zigzag pattern similar to that of <i>Staphylococcus saprophyticus</i> .

TABLE II
NUMBER OF IMAGES REMOVED DURING DATA PREPARATION FOR EACH BACTERIAL STRAIN

Strain of Bacteria	Raw Data	After Manual Removal	After Automatic Mask Identification
<i>Candida albicans</i>	358	316	216
<i>Enterococcus faecalis</i>	362	294	110
<i>Escherichia Coli</i>	806	676	648
<i>Klebsiella pneumoniae</i>	421	308	244
<i>Pseudomonas aeruginosa</i>	362	261	127
<i>Staphylococcus aureus</i>	360	349	212
<i>Staphylococcus saprophyticus</i>	359	356	234
<i>Staphylococcus epidermidis</i>	311	199	116
<i>Streptococcus agalactiae</i>	360	275	99

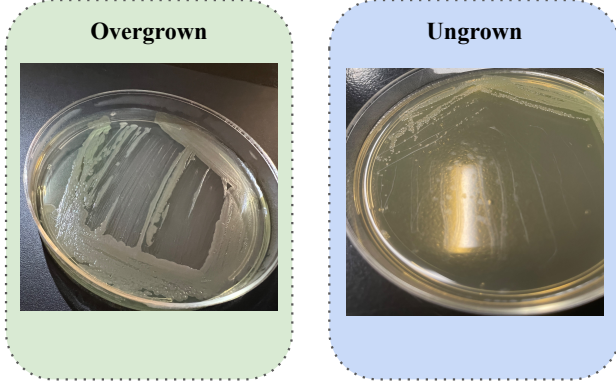


Fig. 2. Example of images removed from the dataset due to overgrowth or minimal growth.

validation set is the remaining data and is 201 images, which we use to measure the performance of the model after training.

C. Model

YOLO, first published in 2015, was created to be an efficient algorithm for object detection [4]. It consists of 24 convolutional layers which precede two fully connected layers. The first 20 layers form the backbone along with an average pooling layer followed by another fully connected layer. YOLO aimed to address object detection as a single-pass regression problem. It does this by dividing the images into an $S \times S$ grid in which each cell (i, j) makes an inference on B bounding boxes, C class probabilities, and a confidence score for each box. This results in a tensor with a shape $S \times S \times (B * 5 + C)$.

For the optimizer, the first YOLO uses stochastic gradient descent. The loss function is shown in (2). In the equation, i refers to the particular cell and j refers to the bounding box predictor. \mathbb{O}_{ij}^{obj} represents if an object appears in cell i and C is the number of classes. X and Y are representative of the coordinate pairs representing the center of the mask area, where w and h are representative of the height and width of these areas. s is indicating the aforementioned $S \times S$ grid and p is representative of the conditional class probabilities.

$$\begin{aligned}
 & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{O}_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
 & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{O}_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{O}_{ij}^{obj} (C_i - \hat{C}_i)^2 \\
 & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{O}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\
 & + \sum_{i=0}^{S^2} \mathbb{O}_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2
 \end{aligned} \tag{2}$$

Our model is the segmentation YOLOv8 variant. YOLOv8 follows largely the same architecture as those that came before it with some additions, as shown in Fig. 5. One addition is the incorporation of the FPN in YOLOv8, which is used to reduce the image resolution and increase the number of feature channels. This allows for the detection of objects at different resolutions. Another addition is the Path Aggregation Network (PAN). The PAN takes features from different levels of the network and combines them using skip connections. Skip connections, rather than process a connection between nodes in a neural network at a higher level, they process the connection at a lower level to avoid overfitting. The combination of these two additions allows for the identification of objects at varying scales and resolutions and instance segmentation.

We train our model in a Python environment running Python 3.11.7 and PyTorch 2.3.0. PyTorch is a Python library for creating neural networks. The GPU we used for training was an NVIDIA GeForce RTX 3090. We set the model to train across 150 epochs and implement early stopping which ends training early if there is no improvement across 20 epochs.

III. RESULTS

We evaluate the model's performance in predicting the nine bacterial strains when shown new images of petri dishes. We choose mAP (mean Average Precision), precision, and recall for performance evaluation of the model. The mAP metric in particular is calculated at an intersection over union (IoU)

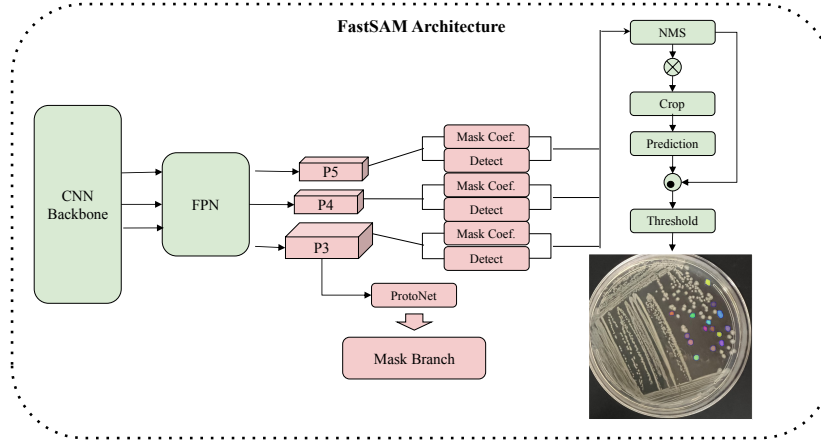


Fig. 3. FastSAM architecture overview and an example image with overlaid masks.

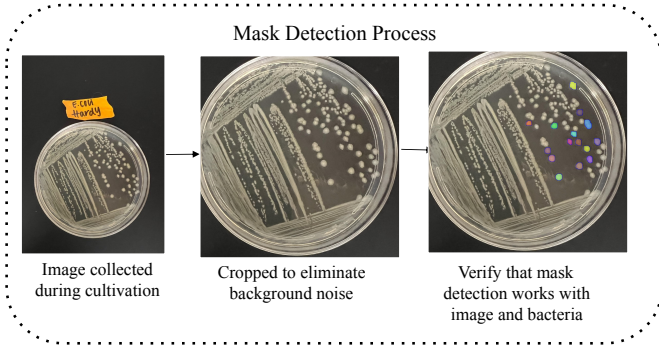


Fig. 4. An example of data preparation including cropping and mask detection.

threshold of 0.50, referred to as mAP50. This means that if 50% or more of the predicted mask overlaps with the ground truth mask, it is considered a correct detection. The mAP metric consolidates precision and recall into a single score [13], calculated as shown in 3 where N is the total number of classes within the dataset and AP_i is a measurement of the precision-recall balance for a specific class:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

Training loss reflects how well the model learns from the data [14]. Validation loss indicates how well the model can identify or classify when shown new data that has not been introduced before [14]. Fig. 6 illustrates the training and validation loss graph, showing an exponential decrease over epochs and converging to a minimum at 119 epochs due to early stopping [15]. The early stopping occurred to avoid overfitting as shown by the trend at epoch 139. The patterns we observe in the graph corroborate the model's capability to learn intricate features and optimize predictions efficiently during training.

Out of the strains predicted by the model to be a certain bacteria, precision is a metric that measures the proportion

of which are predicted to be true. Equation (4) portrays precision, which is the ratio of true positives (TP) over all of the bacteria strains the model predicts correctly ($TP + FP$) [16]. Our model's precision value averages 0.756 with a standard deviation of 0.111. Precision measures the proportion of accurate results, predicted by the model, to serve as a useful indicator to identify bacteria for medical labs with limited resources. Out of all the actual bacteria with the correct labels, recall measures the proportion that the model finds correctly. Equation (5) demonstrates recall, which measures the amount of (TP) over all of the bacteria strains that are accurately predicted by the model ($TP + FN$). Recall serves to indicate the accuracy of the model in identifying similar data outside of training data [16]. Our model's recall value averages 0.71 with a standard deviation of 0.089. Fig. 8 illustrates the high recall and precision obtained from the model which ensures high accuracy of detecting as much of a bacteria as possible (recall) and detecting as many correctly identified bacteria as possible (precision). In the figure, precision has the most positive variation in metric value in the initial 50 epochs, which indicates the model's focus on ensuring as much correct bacteria is identified in the initial stages of training. Then the precision curve changes similarly to the recall curve in the later epochs which shows more focus on ensuring high identification with identified bacteria regardless if it is correctly matched (recall) and bacteria that are correctly identified (precision). The mAP50 in Fig. 8 represents the combination of precision and recall. Our model's mAP50 score averages 0.791 and has a standard deviation of 0.119. The curve tends to increase below the precision and recall and then consistently produces metric values above the precision and recall curves at around 70 epochs. This suggests that in the proceeding epochs, it tends to maximize both precision and recall which is consistent with the analyses found in the individual precision and recall curves.

$$\text{Precision} = \frac{(TP)}{(TP) + (FP)} \quad (4)$$

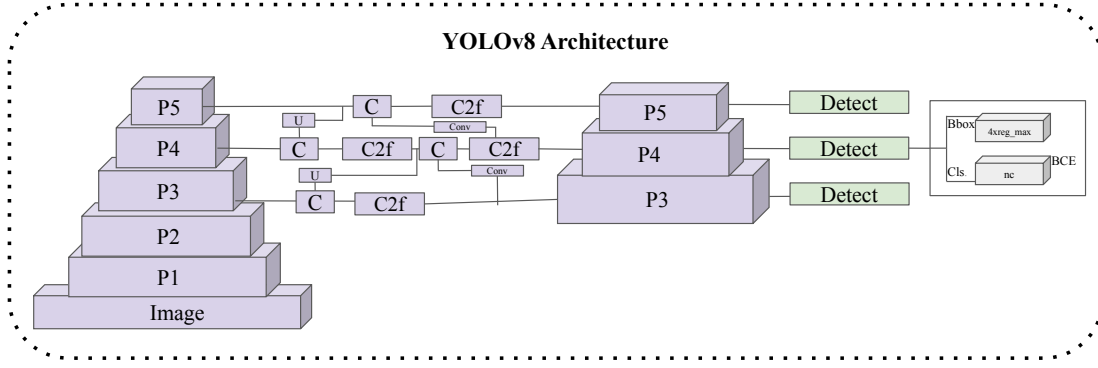


Fig. 5. YOLOv8 architecture overview



Fig. 6. The training and validation loss graph demonstrates how well the model learns and its ability to generalize to new data.

$$\text{Recall} = \frac{(TP)}{(TP) + (FN)} \quad (5)$$

A confusion matrix evaluates the performance of a classification model by showing the counts of actual versus predicted classification of each class. In the confusion matrix in Fig. 9, the darker shades of blue represent greater accuracy of the model to predict the label of certain bacteria strains. The diagonal of the matrix illustrates the number of points for which the predicted label is true [17]. The diagonal values present on our confusion matrix range between 1 and 0.87. The bacteria strains were all accurately identified by the model with a high percentage at and above 87%. The percentage of incorrect predictions of the model is at most 9%. The model is able to predict correct bacteria strains at a high, statistically significant amount of accuracy to provide a reliable model for the identification of bacteria strains in patients.

Fig. 7 shows the precision-recall curves for each bacteria type and a combined curve representative of all the classes in blue. *Enterococcus faecalis* in orange is shown to have the lowest precision-recall curve which indicates the models struggle in correctly identifying this bacteria due to a lack of valid data as shown in II. *Staphylococcus aureus* has the highest trade-off and curve in the graph, which indicates how the model can most accurately identify this bacteria type. The blue curve representative of all the classes portrays a high precision-recall which reflects the model ability to predict the correct bacteria type to its designated label.

By analyzing the performance of each metric, we affirm the model's ability to generalize well to unseen data and maintain high accuracy in real-world applications. The model's performance metrics, including the mAP50 score, precision, recall, and normalized confusion matrix, substantiate its efficacy and promise for further advancement instance segmentation has in the field of bacterial classification.

IV. DISCUSSION

The findings in this study highlight the potential of using advanced computer vision models, specifically instance segmentation, for the identification of bacterial strains in urine samples. The combination of the YOLOv8 model with FastSAM for mask generation has demonstrated a high degree of accuracy in identifying and segmenting bacterial colonies. This is evident in our model achieving an average mAP50 of 0.79. This level of performance highlights our model's capability to effectively localize and classify bacterial strains, making it a promising tool for diagnostic applications. This is a tool that is particularly useful in resource-limited medical institutions given its low computational need [9].

Despite the promising outcomes, there are areas for improvement. The current model is efficient with nine strains of bacteria, thus the next steps would be to expand its classification capabilities to include more strains. This extension will require the collection of additional data to enhance the model's database and knowledge. Moreover, further research should include the manual labeling of each image to provide the model with more instances for training. To achieve this,

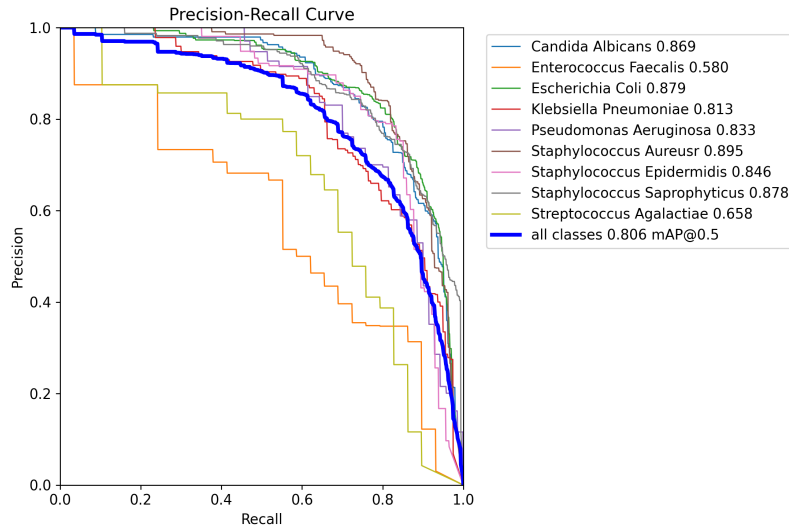


Fig. 7. The trade-off between precision and recall for each bacteria type and a combined curve for all classes.

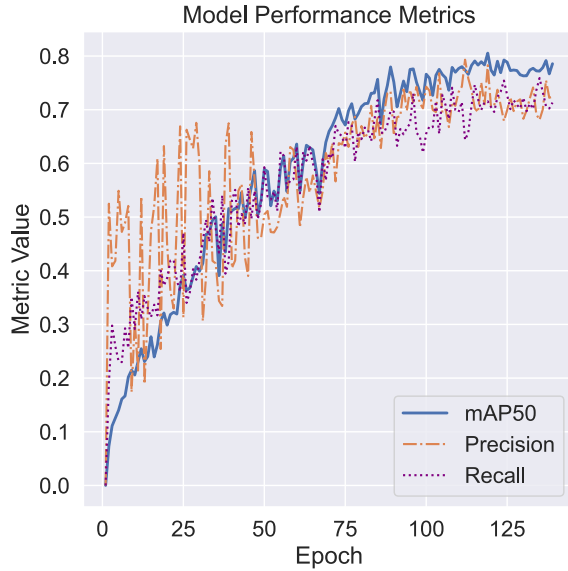


Fig. 8. The model performance metrics graph portrays the amount of precision, recall, and mAP50 the model has tracked across 139 epochs.

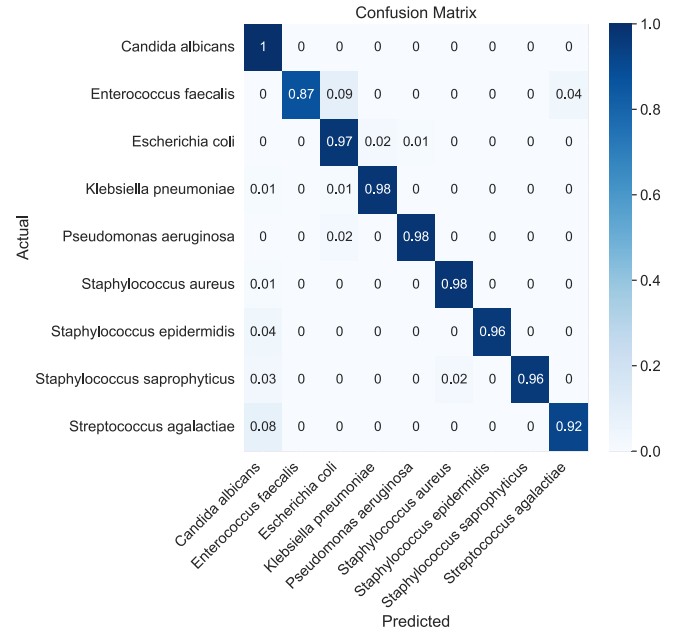


Fig. 9. Confusion matrix to demonstrate the accuracy of the model to correctly predict bacteria strains

more data would need to be collected to add to the model's current database and knowledge.

Additionally, future developments may include creating a mobile application to make this model more accessible. It would allow healthcare professionals to take pictures of petri dishes for analysis, reducing the time needed to accurately identify the bacteria affecting a patient and improve treatment outcomes.

V. CONCLUSION

In this report, we introduced an innovative approach for the identification of bacterial strains using an instance segmenta-

tion model. Using the FastSAM technique for mask generation to train a YOLOv8 model, we developed a system capable of detecting and classifying bacterial colonies among nine strains. The model achieved a notable mAP50 score of 0.79 and demonstrated strong performance as seen in the normalized confusion matrix, indicating its potential as a diagnostic tool in resource-limited settings.

Future research will focus on enhancing the dataset with additional bacterial strains and detailed manual annotations to further improve the model's performance. This will involve maximizing the overall performance of our model before pro-

ceeding with real-world testing. By addressing these areas, our goal is to refine the model and contribute to the advancement of bacterial detection and patient treatment.

REFERENCES

- [1] Xiaorong Yang, Hui Chen, Yue Zheng, Sifeng Qu, Hao Wang, and Fan Yi. Disease burden and long-term trends of urinary tract infections: A worldwide report. *Frontiers in public health*, 10:888205, 2022.
- [2] Ana L. Flores-Mireles, Jennifer N. Walker, Michael Caparon, and Scott J. Hultgren. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nature reviews. Microbiology*, 13:269–84, May 2015.
- [3] Aditya Vasan and James Friend. Medical devices for low- and middle-income countries: A review and directions for development. *Journal of medical devices*, 14:010803, Mar 2020.
- [4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016.
- [5] Mupparaju Sohan, Thotakura Sai Ram, and Ch. Venkata Rami Reddy. A review on yolov8 and its advancements. In I. Jeena Jacob, Selwyn Piramuthu, and Przemyslaw Falkowski-Gilski, editors, *Data Intelligence and Cognitive Informatics*, pages 529–545, Singapore, 2024. Springer Nature Singapore.
- [6] Xu Zhao, Wenchao Ding, Yongqi An, Yinglong Du, Tao Yu, Min Li, Ming Tang, and Jinqiao Wang. Fast segment anything, 2023.
- [7] Adam W. Harley, Konstantinos G. Derpanis, and Iasonas Kokkinos. Segmentation-aware convolutional networks using local attention masks. *CoRR*, abs/1708.04607, 2017.
- [8] Paolo Andreini, Simone Bonechi, Monica Bianchini, Alessandro Mecocci, and Vincenzo Di Massa. Automatic image analysis and classification for urinary bacteria infection screening. In Vittorio Murino and Enrico Puppo, editors, *Image Analysis and Processing — ICIAP 2015*, pages 635–646, Cham, 2015. Springer International Publishing.
- [9] Vinod Kumar Chauhan, Kalpana Dahiya, and Anuj Sharma. Problem formulations and solvers in linear svm: a review. *Artificial Intelligence Review*, 52(2):803–855, 2019.
- [10] Hongda Wang, Hatice Ceylan Koydemir, Yunzhe Qiu, Bijie Bai, Yibo Zhang, Yiyin Jin, Sabiha Tok, Enis Cagatay Yilmaz, Esin Gumustekin, Yair Rivenson, and Aydogan Ozcan. Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning. *Light: Science & Applications*, 9(1):118, 2020.
- [11] Donald Breakwell, Christopher Woolverton, Bryan MacDonald, Kyle Smith, and Richard Robison. Colony morphology protocol. Technical report, American Society for Biology, 2007.
- [12] Jiayi Fan, JangHyeon Lee, InSu Jung, and YongKeun Lee. Improvement of object detection based on faster r-cnn and yolo. In *2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, pages 1–4, 2021.
- [13] Mehmet Şirin Gündüz and Gültekin Işık. A new yolo-based method for real-time crowd detection from video and performance analysis of yolo models. *Journal of Real-Time Image Processing*, 20(1):5, 2023.
- [14] Samet Oymak, Mingchen Li, and Mahdi Soltanolkotabi. Generalization guarantees for neural architecture search with train-validation split. In *International Conference on Machine Learning*, pages 8291–8301. PMLR, 2021.
- [15] Manuel Vilares Ferro, Yerai Doval Mosquera, Francisco J. Ribadas Pena, and Víctor M. Darriba Bilbao. Early stopping by correlating online indicators in neural networks. *Neural Networks*, 159:109–124, 2023.
- [16] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd International Conference on Machine Learning, ICML '06*, page 233–240, New York, NY, USA, 2006. Association for Computing Machinery.
- [17] Jasmina Novakovic, Alempije Veljovic, Sinivsa Ilic, Željko M. Papic, and Tomovic Milica. Evaluation of classification models in machine learning. *Theory and Applications of Mathematics & Computer Science*, 7:39–46, 2017.