



Harnessing deep learning for faster water quality assessment: identifying bacterial contaminants in real time

Fahad Ahmed Khokhar¹ · Jamal Hussain Shah² · Rabia Saleem³ · Anum Masood⁴

Accepted: 23 March 2024
© The Author(s) 2024

Abstract

Water is essential for human survival. Humans can live without food for a few days but without water, a person can barely survive for 3–5 days. Various parts of the world, particularly under-developed countries, have regions where clean water is scarce, and humans living in such conditions have no access to clean water. Our solution provides information on whether water is contaminated or not. Moreover, it overcomes the delay time in getting the result of water contamination using traditional methods of up to 5–6 hrs. Our proposed method detects the colonies of the bacteria that are taken from the water sample (after gram staining) and then classifies the type of bacteria to whom it belongs and how much quantity of each bacterium causes infection to the human body. Bacteria detection is performed by a novel deep learning-based model with user-specified parameters. To improve our ability to detect dangerous bacteria including *E. coli*, yeast, and particles, we perform tests using datasets from a variety of researchers. On the test benchmark, the fine-tuned proposed model achieves 84.56% accuracy and provides the level of contamination in water.

Keywords Optimized Fast Detection · Water Contamination Assessment · YOLOv5 · Fine-Tuning · Rapid Detection · Bacterial Detection · Deep learning

1 Introduction

With the rapidly growing population, it is predetermined that the world population will increase by about 9 billion in the

year 2050 [25]. From this perspective, the most critical issue will be access to clean water for drinking. Moreover, the absence of clean water will be a great concern for disaster-related human health [5]. According to the World Health Organization (WHO), almost 785 million people have access to clean water whereas 2 billion people are deprived of clean water and are feeble to drink contaminated water [21]. To ensure the purification of water for safe consumption, there is a need for effective methods to eliminate bacteria, viruses, and toxic pollutants like heavy metal ions [1, 2]. Contaminated water is responsible for causing numerous harmful diseases, resulting in over 5 million deaths worldwide each year [22]. Bacterial pathogens in drinking water, such as *Escherichia Coli*, *Legionella*, and *Campylobacter Jejuni*, are major contributors to human mortality, causing diseases like cholera, cryptosporidiosis, giardiasis, and gastroenteritis [8]. Additionally, antimicrobial resistance is a pressing global health concern, impacting the effectiveness of antibacterial compounds in combating infectious diseases caused by resistant microorganisms [13].

The research in object detection for rapid bacteria identification in Gram stain images using deep learning lacks specialized models and methodologies for water samples.

✉ Anum Masood
anum.masood@ntnu.no

Fahad Ahmed Khokhar
fahadahmed.khokhar@unifi.it

Jamal Hussain Shah
jshah@ciitwah.edu.pk

Rabia Saleem
rabia_278@hotmail.com

¹ Department of Mathematics and Information, University of Florence, Florence, Italy

² Department of Computer Science, COMSATS University Islamabad, Wah Cantt, Pakistan

³ Department of Computer Science, Government College University, Faisalabad, Pakistan

⁴ Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

Current methods are time-consuming, but deep learning models like YOLOv5 offer real-time capabilities. There's a need for accurate models that can detect and classify bacteria types, quantifying their quantities to assess water contamination and health risks. The fine-tuning process for such models is underexplored. Developing a specialized deep learning model can revolutionize water quality assessment, improving monitoring systems and public health outcomes by mitigating risks from contaminated drinking water.

The presence of *E. coli* and Total Coliform bacteria in the drinking water is the main cause of contamination of the water and the source of the most common disease which is waterborne pathogens [27]. According to Environmental Protection Agency (EPA) guidelines, traditional methods including culturing samples gathered using solid agar plates or liquid media are used to detect and recognize *E. coli* and Total Coliform under the supervision of bio researchers through counting the numbers [7, 31]. The liquid growth media is used to identify fecal coliform bacteria with high specificity and hence requires at least 18 h for full growth to study. Solid agar plates are a cheaper method that allows for greater discretion in the number of samples to be tested, which can range from 100 mL across many liters to improve sensitivity. However, the traditional detection of bacterial method needs almost 24–48 h for the growth of the bacteria to make the colonies.

Unlike conventional detection methods that solely focus on the presence of bacteria, our proposed approach exhibits a transformative feature: classification of bacterial types and contamination levels. This innovative aspect enables the system to discern specific dangerous bacterial pathogens, such as *E. coli*, yeast, and other particles, known to cause severe infections in the human body. The basis of our approach lies in the employment of the fine-tuned YOLOv5 (You Only Look Once version 5) model, notable for its real-time capabilities and dense prediction. YOLOv5 is a single-stage object detection model that has been fastidiously adapted for the particular task of identifying bacterial colonies in Gram stain images. Our novel implementation is accomplished through customization, incorporating additional convolutional layers after the backbone and presenting two large-scale convolutional layers in the line. These architectural alterations appoint the model with enhanced feature extraction and differentiation capabilities, efficaciously reducing repetitiveness and rising semantic feature extraction. The result is a powerful model capable of achieving higher recognition accuracy and accommodating seamlessly to diverse scenarios, crucial for precise bacterial detection.

Additionally, the model quantifies the quantity of each bacteria, furnishing valuable information about the extent of contamination in the water. This comprehensive analysis equips decision-makers with critical insights, facilitat-

ing targeted actions to combat waterborne threats effectively.

In the subsequent sections, we will present details of the proposed technique. Section 2 will provide comprehensive related work. Section 3 presents an overview of the methodology, including the fine-tuning process of YOLOv5, the incorporation of three additional layers, and the utilization of the K-means++ algorithm for anchor selection. Section 4 and Sect. 5 explain dataset preparation and preprocessing details. Sect. 6 will present the experimental setup and evaluation, showcasing the dataset used, performance metrics, and comparison with traditional methods. Finally, we will conclude the paper in Sect. 7, summarizing the key findings and emphasizing the significant contribution of this research.

The key contributions of this paper can be summarized as:

1. This research enhances the YOLOv5 model to detect bacteria in Gram stain images by adding extra Convolutional layers, particularly large-scale ones in the Header, improving object classification and identification of harmful bacteria in water samples.
2. The study integrates three extra Convolutional layers post-Backbone and includes two large-scale convolutional layers in the Header, enhancing feature extraction and differentiation of the YOLOv5 model. These adjustments yield more distinct features, diminishing redundancy and bolstering semantic feature extraction, ultimately enhancing recognition accuracy crucial for precise bacterial detection across diverse scenarios.
3. The proposed method extends beyond bacterial detection to classify bacterial types in water samples, including hazardous pathogens like *E. coli* and yeast, offering insights into potential infections. Moreover, the model quantifies bacterial quantities, providing valuable data on water contamination levels.
4. The deep learning model in this study achieves an impressive 84.56% accuracy on test benchmarks, ensuring reliable identification of harmful bacteria in water samples. Operating in real time, the system enables swift responses to contamination issues, thereby mitigating risks and enhancing public health by rapidly and accurately detecting dangerous bacteria in drinking water.

In summary, the paper makes three primary advancements: firstly, it employs an efficient and real-time method for detecting bacteria in Gram stain images; secondly, it demonstrates proficiency in classifying bacterial types and assessing contamination levels; and thirdly, it exhibits a high accuracy in identifying harmful pathogens. The suggested approach tackles the crucial concern of ensuring clean water availability for human survival and holds substantial promise for enhancing water quality monitoring systems on a global scale.

2 Related work

This section presents the research goal and establishes the study's theoretical foundation. Numerous state-of-the-art methods and contributions of different researchers are discussed considering the bacterial effect on water which is harmful to human health. Moreover, it describes the role of machine learning in the field of microbiology in the detection and recognition of multiple pathogens and how it overcomes the issues in the past decades.

2.1 Traditional methods

Traditional methods for bacterial detection, such as the widely utilized culturing on agar plates, have long been employed; however, they suffer from significant time constraints, taking as much as 24–48 h to yield results [19]. This delay poses a considerable risk in situations where swift detection is crucial to prevent waterborne diseases and ensure public health safety.

The culture method in combination with the 'negative to date' concept [18] is effective but does not have any short period for detection. Chain Reaction methods including PCR and Q-PCR are quite useful for rapid bacterial detection [9, 20, 26]. Rapid detection also may use the culture method for getting the sample [18] ready for the rapid detection analysis technique. This kind of culture method needs 7 days of incubation to make the sample analyzable for rapid detection work.

2.2 Machine learning methods

There are multiple methods to detect the bacteria from the images including the statistical method [29], and Artificial Intelligence (AI) based methods [28]. According to [28], an algorithm uses geometric properties such as curvature, density, irregularity, vascularity, and length-to-width ratio to identify bacterium species. Furthermore, because the bacillus' shape is not a distinguishing factor (due to the based architecture in various microbial species), this color is used as a salient approach.

Supervised learning methods

Men et al. [14] used a support vector machine (SVM) to recognize heterotrophic bacteria colonies in images. Denoising, luminance balancing, softening, and enhancing were applied to the obtained colored bacterial colony picture. The picture was then segmented using an appropriate threshold approach. Color consistency between colonies was employed for chromatic pictures. SVM was then trained to categorize and count extracted colonies based on the color and shape attributes of bacterial colonies. Xiaojuan et al. [32] suggested

a machine learning-based technique for recognizing wastewater bacteria species in images. Following that, morphology and constant moment-based characteristics were recovered, and dimensionality was reduced using Principal Component Analysis (PCA).

Unsupervised learning methods

Chayadevi et al. [4] used unsupervised learning methods to extract microbiological clusters from microscopic pictures. There were 320 digitized microscopic photographs of bacterium species in the dataset. Image preprocessing was done to use the threshold approach and pattern classification. Then, with 81 characteristics such as perimeter, eccentricity, and circularity, a feature set was created. After that, using an ANN approach called self-organizing map (SOM) and the K-means clustering technique, bacterial clusters were identified.

2.3 Deep learning methods

The emergence of deep learning models has shown promising potential in various object detection tasks. Deep Learning (DL) classifiers [15] are also used for bacterial classification. Single-stage object detection techniques like YOLOv5 have particularly garnered attention due to their real-time capabilities [34]. This opens up new avenues for rapid bacterial detection in Gram stain images, significantly reducing the time required for assessment compared to traditional methods. The ability of deep learning models to handle complex patterns and extract high-level features from images has led to their consideration as a suitable approach for this application [17].

However, one critical limitation lies in the scarcity of specialized deep learning models and methodologies explicitly tailored for bacterial detection in water samples [24]. While [30] proposed a deep learning model in which new bacteria classification based on Big Transfer (BiT) that mutually with weight initialization based-rectified linear unit (WIB-Relu) activation and graph Laplacian-based data cleaning. This method reached outcomes with an accuracy of 99.11%, precision of 99.31%, recall of 99.09%, and F1 score of 99.06% and tested on microscopic bacteria images of a public dataset called Digital Images of Bacteria Species (DIBaS)

While general deep learning models have shown success in other domains, their application to waterborne bacteria detection requires fine-tuning and optimization to adapt to the unique characteristics of bacterial colonies [12].

In regard to the identification of the TB bacteria within ZN slide pictures, Osman et al. [23] used a mix of genetic algorithms (GA) and artificial neural networks (ANN). A total of 960 TB pictures were collected from 120 slide photographs in the collection. Image segmentation employing

the color segmentation method, K-means clustering, as well as a region-expanding algorithm were all part of the suggested method. After that, a median filter was used to remove the noise. Histograms were retrieved as features after that. GA was used to choose the best characteristics

Xu et al. [33] introduced a methodology to categorize images of red tide algae. This method involved the utilization of the Otsu self-adaptive technique to extract the region of interest, followed by the application of the Canny edge detection approach to reduce closed contours. Subsequently, bounding boxes were obtained, and the classification process was executed through the ensemble of Support Vector Machines (SVM), the summation of negative probabilities, and semi-supervised fuzzy C-means clustering. In parallel research, Mosleh et al. [16] delved into the identification of freshwater algae species in images. The dataset encompassed images from five different algae genera: *Oscillatoria*, *Navicula*, *Chroococcus*, *Microcystis*, and *Scenedesmus*.

In recent studies, in [6] the architecture of the Siamese network involves the two convolutional neural networks (CNNs) that share the weights and machine learning algorithm utilized to identify bacteria tested on Raman spectral datasets. In the comparison, the method was evaluated based on training time, prediction time, mean sensitivity, and the number of parameters. The largest mean sensitivity of 83.61 ± 4.73 , data scenarios limited, and handling unbalanced performance outstanding which is achieving the accuracy of 73% prediction.

In [11], the authors proposed an improved method in the IoT network of Bacterial Foraging Optimization with optimum deep learning for Anomaly Detection (IBFO-ODLAD). The results of this method were validated on two datasets of UNSW NB-15 and UCI SECOM which improved performance accuracy by 98.89% and 98.66% of the given IBFO-ODLAD algorithm.

Moreover, the literature highlights the need for deep learning models that can not only detect the presence of bacteria but also classify specific types accurately and quantify their quantities [3]. This level of specificity is essential for precisely assessing contamination levels and potential health risks associated with different types of bacteria in water sources.

3 Methodology

The proposed method presents an innovative technique for optimized and fast detection for accelerated water contamination assessment using fine-tuned YOLOv5 with three additional layers and K-means++ anchor selection. The approach overcomes the delay issues in traditional methods, offering real-time detection capabilities. By enhancing feature extraction and anchor box selection, the method achieves

faster and more accurate detection of waterborne contaminants, reducing the detection time to a fraction of what traditional methods require. The proposed model has diverse real-world applications, including real-time water quality monitoring, early disease outbreak detection, environmental conservation, emergency response, and quality assurance in food production. With its advanced object detection and classification features, the model enables quick identification and quantification of harmful bacteria in water samples, supporting proactive interventions for safe drinking water. It also aids in emergency scenarios, industry quality control, and contributes valuable insights to environmental and microbiological research. The model's integration into practical solutions highlights its role in leveraging artificial intelligence for public health, environmental sustainability, and industrial practices.

3.1 Fine-tuning YOLOv5

For the detection of the objects, we use the Fine-Tuned model of DNN such as Yolo as the proposed model. In YOLO we tuned the architecture by modifying architecture in multiple stages. Single Stage Object Detection like YOLOv5 is used for the implementation of this research work and is comprised of three main components namely Backbone, Bottleneck, and Header to create the dense prediction. Backbone extracts the rich features from the input image and reduces the spatial resolution of an image by increasing its feature in terms of channel resolution. Backbone uses the *CSPDarknet* as the backend platform. The bottleneck is further utilized for the extraction of feature pyramids. It ensures the generalization of objects on multiple sizes and scales. Moreover, the header is used in the output to apply the multiple anchor boxes on feature maps and produce the final bounding boxes along with labels.

In the fine-tuned YOLOv5 model, we add three conv layers to the basic YOLOv5 network architecture after the backbone. As images are very small and difficult to extract features, Convolution Layers play an important role because this conv layer handles the spatial redundancy by sharing their weights. These three additional layers make the retrieved characteristics more unique and informative while also cutting down on unnecessary repetition. This is mainly because of the subsampling layers and the repeated cascaded convolutions. The network presents a compact feature about the image's content by eliminating unnecessary data.

3.2 Enhanced object detection architecture

Deeper convolutional layers in a network allow for better semantic feature extraction, leading to higher recognition accuracy and an easier time adapting to new circumstances. Our proposed model is illustrated in Fig. 1.

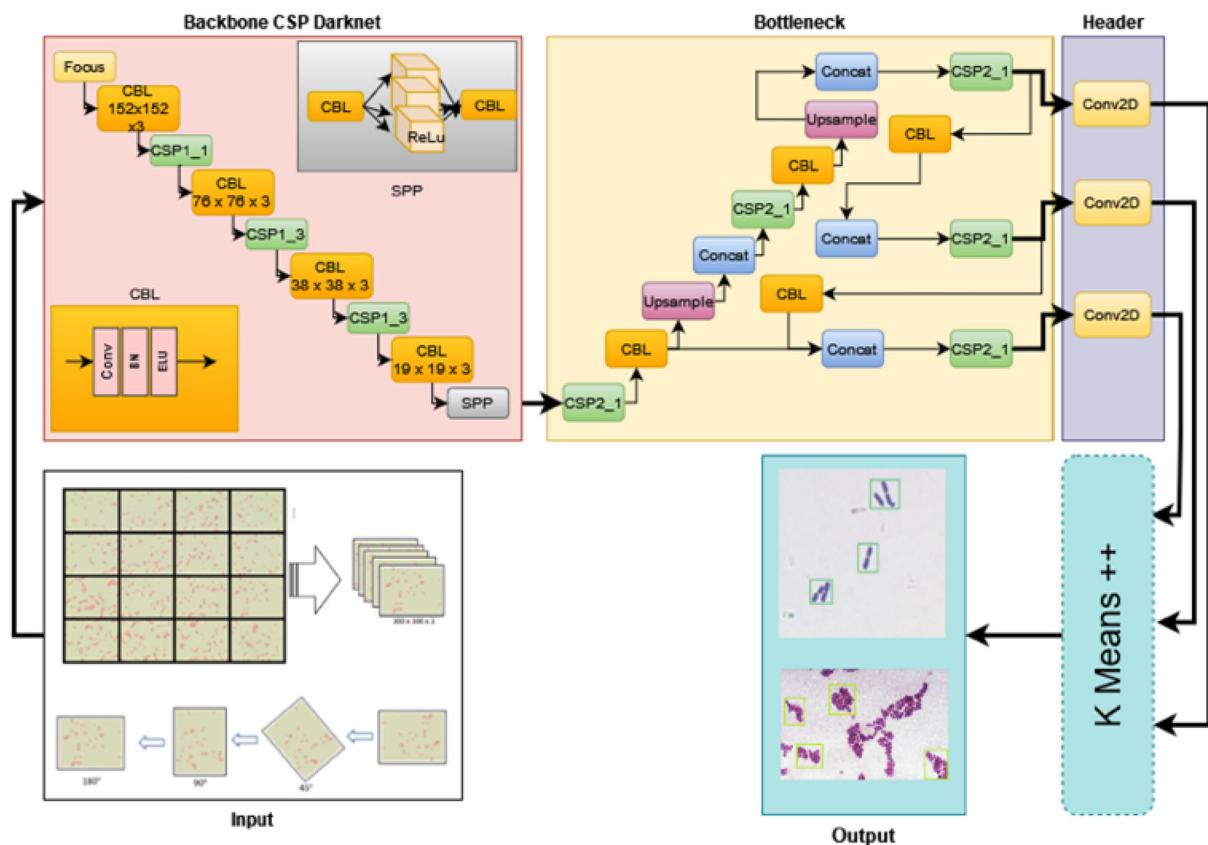


Fig. 1 An Overview of our Proposed Model for Bacterial Types and Contamination Level Classification

The header also received two additional large-scale convolutional layers. Here, we employed a 3×3 and a 1×1 convolutional layer. By using a convolutional layer with a size of 1×1 , the network’s ultimate depth and computation load were both decreased. With this adjustment, the improved network is better able to distinguish between big and tiny, long-range objects. For training the model, we use 70% of the dataset whereas 30% is used for the testing purpose. At the end of the network where the output of the image is shown, we use the K-means++ algorithm to optimize the anchor box detection.

In another model for object detection, we used an SSD network. Faster R-CNN makes limit boxes utilizing a neural network approach and afterward utilizes those cases to classify things. The whole interaction works at 7 frame rates, which is respected best in class as far as accuracy. Undeniably not as much as what genuine handling requires. SSD executes various improvements, including the number of co-highlights and standard boxes, to compensate for the decrease in exactness. These upgrades permit SSD to accomplish the exactness of the Faster R-CNN using lower-resolution pictures, speeding up considerably further. The SSD comprises two parts one is for the extraction of features and

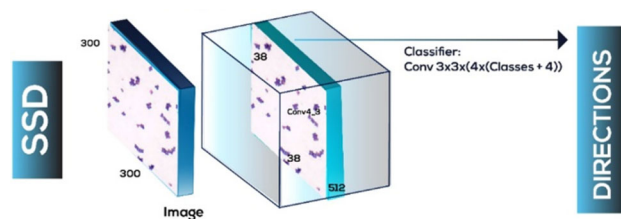


Fig. 2 Overview of the SSD-Net Model with two components (a) Feature Extraction, and (b) Enhanced Object Detection

another one is for applying convolutional filters for object detection.

3.3 Bacterial types and contamination level classification

We use our custom dataset for the extraction of feature maps. Then, it detects the object using the *Conv4_3* layer. Each prediction which layer predicts the composed boundary box and gives a score of 21 to each class and we pick the highest score for the bounded object. The convoluted layer predicts more than 5 boxes in a single bacteria class then we get only one box that has more prediction accuracy as shown in Fig. 2.

4 Experiments

4.1 Ethical approval

The study subject was approved by the ethical review committee (ERC) of the University of Wah, Pakistan. All the bacterial cultures were handled and tests were conducted in compliance with the University of Wah, Pakistan's health and safety regulations.

4.2 Bacterial cultures

E. coli (Migula) Castellani and Chalmers (ATCC® 25922™), Yeast (*Saccharomyces cerevisiae*) Meyen ex E.C. Hansen (MYA-4941™), and Particles are used as our culture organisms.

4.3 Preparation of gram stain slides

Bacterial images were produced by Gram staining and light microscopy of the freshly grown cultures. The slides were prepared by adding a small amount (with the help of a wire loop) of culture and suspending the bacterial cells in a drop of distilled water. Bacteria were heat-fixed on the glass slide after drying the smear. Crystal violet was added to the smear and washed with water after 1 min. Gram's Iodine was added and washed with water after 1 min and then rinsed with ethanol solution (70%). Then the slide was washed immediately with plenty of water to remove excessive amounts of ethanol. Safranin was added to the smear and washed with water after 1 min. The slides were air-dried and a drop of immersion oil was added to the smear to get ready for further processing.

5 Dataset preprocessing

The images of bacteria were observed at 100X magnification by using camera-fitted Light microscope (IM910, IRMECO Germany). Image preprocessing steps are used to make the image sharp and clearer. The bacterial images are blurry because of the image splitting and the detail of the image is not shown to be clear. Due to microscopy images of the slides, the image should be clear and have high detail for further processing, i.e., training.

The dataset contains 3 categories, i.e., *E. coli*, Yeast, and Particle with almost 20 images in each dataset comprising of $2048 \times 1532 \times 3$ size. The image size is too huge for processing the training due to which the image is split into $300 \times 300 \times 3$ chunks for making the dataset increase in

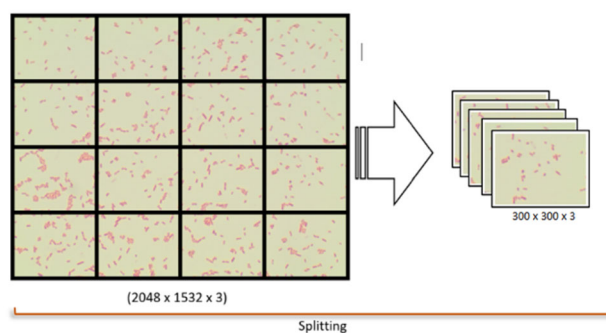


Fig. 3 Pre-processing for image enhancement of the bacterial gram stain slides

numbers as well as suitable for the training of the model as shown in Fig. 3.

5.1 Image translation

After splitting the images dataset increases to 200 images in each class. Due to the low number of images, the images are further augmented using the data augmentation techniques, i.e., Image Rotation, Image Flipping, and Image Translation are used to maximize the no of images given in Eq. 1. The dataset increases from 200 to 998 images per class.

$$\begin{bmatrix} I_x \\ I_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & r_x \\ 0 & 1 & r_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I'_x \\ I'_y \\ 1 \end{bmatrix} \quad (1)$$

5.2 Image flipping

In this, we flip the image to make the increase in the number of images. Different flipping techniques are involved in data augmentation including Horizontal and Vertical flipping. In the horizontal flipping, reflection is consistently over the y-axis mentioned in Eq. 2 also vertical flipping is completed through reflection against the x-axis. The pixel present in the coordinate of x, y coordinate will become (width - x - 1, y) in the output image with new origination.

$$I_{(x,y)} = I_{(\text{width}-x-1,y)} \quad (2)$$

5.3 Image rotation

The image rotation technique is used for the augmentation of the dataset. In the image rotation, images are rotated in 45° , 90° , and 180° . To compute the new image, we make iterations throughout the image pixel by pixel and print the corresponding pixel from the given source image as shown

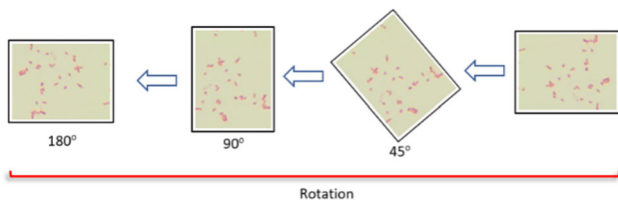


Fig. 4 Affine transformation (rotation) using multiple angles is used as a data augmentation technique

in Eq. 3. The rotation of the image is shown in Fig. 4.

$$\begin{bmatrix} I_x \\ I_y \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I'_x \\ I'_y \\ 1 \end{bmatrix} \quad (3)$$

In Eq. 4, δ represents the output of the trained network and I is the input vector that corresponds to the target input image. We changed the SSD slightly bit by getting work from only the *Conv4_3* layer because our problem only focuses on the three classes of bacteria and each microscopic image has less than 3 or 4 colonies of bacteria. We did not add extra layers to make the complexity of the model by getting the work using a convoluted layer whose input is $300 \times 300 \times 3$ and then further it convoluted with the $38 \times 38 \times 512$ features.

$$SSD(\delta, I) = \sum_i \left(\frac{I_i}{\|I\|_2} + \frac{1 - I_i}{\|1 - I\|_2} \right) (\delta_i - I_i)^2 \quad (4)$$

5.4 Implementation details

Nvidia GeForce RTX2060 GPU hardware was used to train the modified YOLOv5 network, using 30,000 iterations. The configuration file was modified as follows to train these two networks:

- A 300×300 pixels picture was used as the input.
- In order to prevent errors caused by insufficient memory, the subdivision and batch parameters were set to 1 and 64, respectively.
- The pace of learning rate was slowed down to 0.005; depending on the number of classes, the step parameter was set to 24,000 or 27,000 (corresponding to 80% and 90% of the total number of repetitions), and the filter size in three convolutional layers next to the YOLO layers was adjusted to 30, from 0 to 1.

CUDA (Compute Unified Device Architecture) Toolkit 10.0, CUDNN (CUDA Deep Neural Network library) 8.2, Visual Studio 2017, and OpenCV (Compute Unified Vision API) 4.0.1 are utilized, among other software and libraries. Figure 5 shows a graph of the number of iterations vs the loss

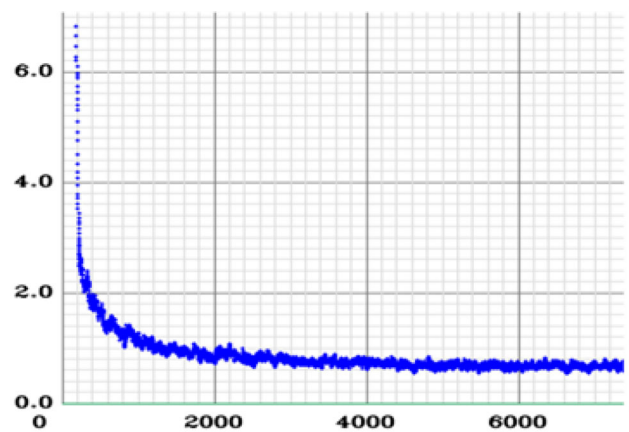


Fig. 5 Loss graph during training of customized YOLOv5

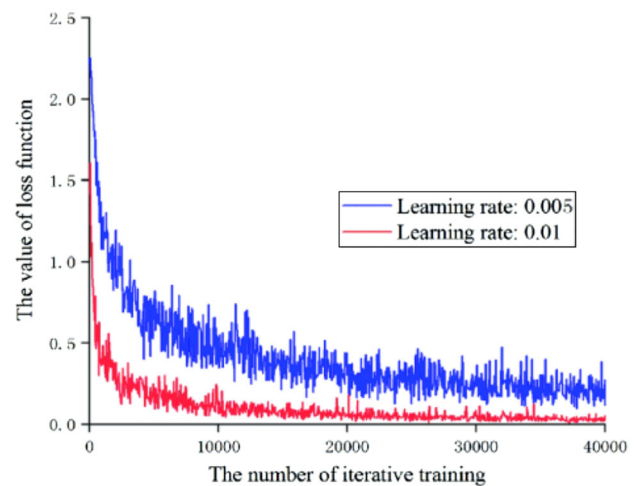


Fig. 6 Graph presenting the training loss over the epochs for SSD-Net

that was generated during training Customized YOLOv5. Figure 6 shows a graph of epochs against the training loss for SSD-Net. The loss for the modified implemented technique was 0.58, after 30k iterations.

6 Results and discussion

Approximately 30% of the entire dataset was used to start the testing phase after the network training was finished. At this point, we were looking at how well the network could identify Bacteria and choose the optimal bounding box that included the object. An area of detection for Bacteria is established by the suggested learning network. The dataset includes multiple Bacteria of varying sizes and forms, necessitating the suggested approach to generate several bounding boxes for each.

However, the non-maximum suppression (NMS) technique is required to pick the best bounding box out of the

Table 1 Performance evaluation of our model modified YoloV5 (M-YoloV5) for classification of E.Coli, yeast, and particles using gram stain slide images

Class	Accuracy	Precision	Recall	F1-Score
E. Coli	0.861	0.681	0.892	0.771
Yeast	0.924	0.972	0.821	0.893
Particles	0.905	0.902	0.831	0.871

Table 2 Performance evaluation of SSD-Net for classification of E.Coli, yeast, and particles using gram stain slide images

Class	Accuracy	Precision	Recall	F1-Score
E. Coli	0.643	0.620	0.470	0.541
Yeast	0.567	0.160	0.260	0.203
Particles	0.785	0.721	0.661	0.691

many possible ones [10]. Using this approach, we were able to narrow down our options to the best possible Bacteria boxes by removing the ones with the lowest confidence scores. Figure 8 suggests that the green box is the most suitable for enclosing a Bacteria; however, the other two bounding boxes also cover some of the bacteria and are therefore options for bacteria detection as well. Based on the algorithm's scoring system, the bounding box with the most confidence is chosen, and then, the other boxes that overlap it the most are discarded. This procedure is repeated until no more representative boxes contain the bacteria.

The non-maximum suppression (NMS) algorithm may be broken down into the following manageable chunks of work:

1. Pick the prediction's margin of error that's the largest;
2. The area of union (IoU) may be determined by solving Eq. 5 which describes the intersection and overlap of the chosen box and the other boxes.

$$IoU = \frac{\text{Area of } \cap}{\text{Area of } \cup} \quad (5)$$

3. Remove any cells that share more than 10% of their IoU with the currently chosen cell.
4. Repeat from steps 1–3.

6.1 Quantitative Results

In order to demonstrate the effectiveness of the proposed method, we evaluated the performance of the proposed model on our in-house dataset. The quantitative results are depicted in Tables 1 and 2, respectively.

6.2 Object detection using proposed model

In YoloV5 after an alteration of the model, we trained the model and got the trained weights for real-time detection and recognition of the bacteria in water. The dataset is first labeled, and each individual image is saved along with its bounding box location. While training the dataset it is found that the precision reaches 48% while the accuracy is about 84.56% as shown in the confusion matrix. Similarly, the mAP for the training of the dataset is about 81% as shown in Fig. 7. Testing results are shown in Table 1 with different parameters for calculation of accuracy, precision, recall, F1-Score, etc., using Fig. 9. The accuracy of the overall model is 84.53% whereas the Kappa is 0.786 as described above.

However, the accuracy of SSD-Net is 49.81% whereas Kappa is 0.247 as shown in Table 3. At SSD-Net we achieve a precision of 0.62 for the E. coli class which is less than the proposed methods. Moreover, there is an unexpected change in the F1 Score that is shown in Table 2. All accuracies of the model are drawn using the Confusion Matrix, which is shown in Fig. 10.

The results and analysis indicate that the proposed adaptation of YOLOv5 outperforms the existing SSD-Net, particularly in terms of accuracy and the Kappa value. The modified YOLOv5 demonstrates higher accuracy, signifying more precise identification of bacteria in water samples. Moreover, the elevated Kappa value suggests a stronger correlation between the model's predictions and the actual ground truth, making the proposed modified YOLOv5 more effective for the specific task of detecting bacteria in water samples. The Kappa value, serving as a metric for assessing agreement between predicted and ground truth values, attains a substantial value of 0.768 for the modified YOLOv5, indicating a notably robust association between the model's predictions and the actual ground truth labels. Conversely, the lower Kappa value of 0.247 for SSD-Net implies a less robust agreement between its predictions and the ground truth.

In comparison to YOLOv5 in terms of purpose and application, the proposed method focuses on the real-time detection of harmful bacteria in water samples using deep learning techniques, specifically employing a DNN model. In contrast, YOLOv5 functions as an object detection model primarily designed for real-time object detection in images and videos. Despite both methods utilizing deep learning, their intended applications and target objects differ significantly. Another distinction between our method and YOLOv5, related to data source and preprocessing, is that our proposed method aims to detect bacteria in Gram stain images derived from water samples. This involves specific preprocessing steps, such as identifying bacterial colonies and classifying the type and quantity of bacteria present. On the other hand, YOLOv5 is typically applied to broader object detec-

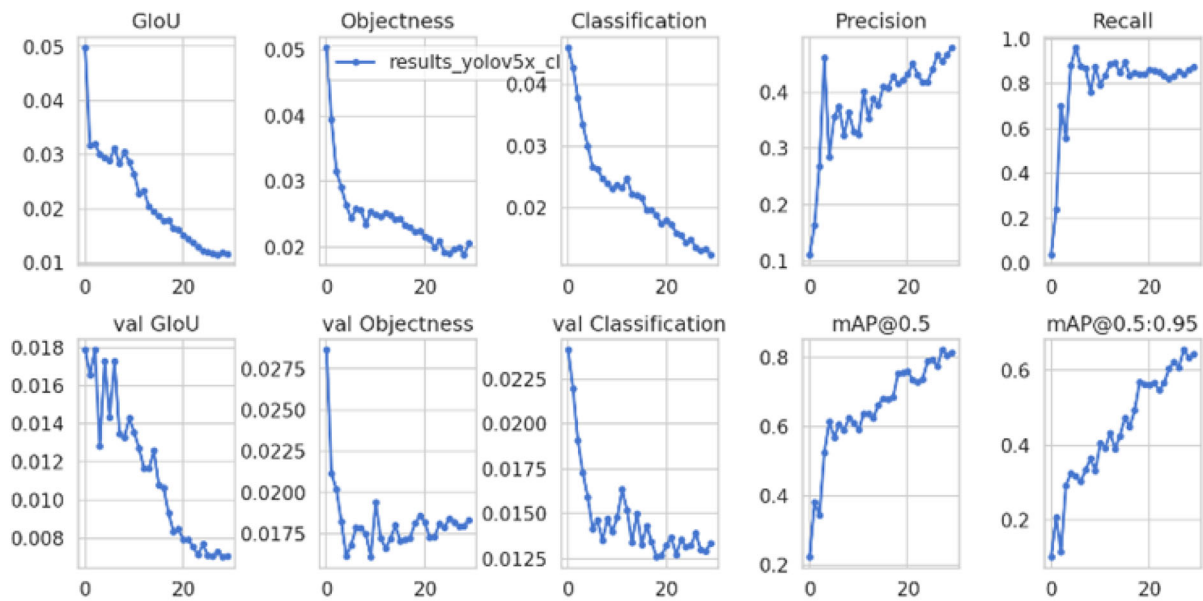


Fig. 7 Visualization of performance evaluation of modified YoloV5 using performance metrics including GIoU, objectness, classification, precision, recall, mean average Precision (mAP)

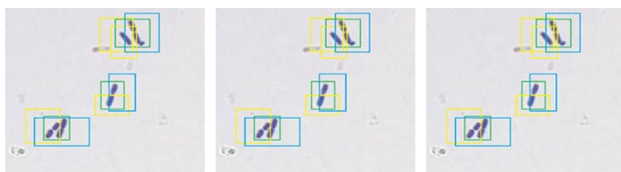


Fig. 8 Box selection method using proposed non-maximum suppression (NMS)

tion tasks and lacks a specific focus on bacterial detection or Gram stain images.

7 Conclusion

Detection of bacteria in different eatables like food and water is a challenging and time-consuming task. An innovative way of testing water contamination using Machine Learning is introduced for earlier detection and recognition of bacteria. Multiple bacteria are found in the water which is considered harmful for human health. We proposed a model that detects and recognizes the water complexity using the microscopic image of the gram stain water. Moreover, the time of detecting the bacteria is reduced from 24h to 12h. Our proposed model YoloV5 is fine-tuned with multiple parameters to get the accuracy up to 84.56% on our customized dataset. The dataset is collected by taking samples from different places and then slides are prepared in a laboratory under suitable temperatures. After gram staining slide is placed inside the microscope and the image is for further processing of the contamination of water. We provide the rapid detection and

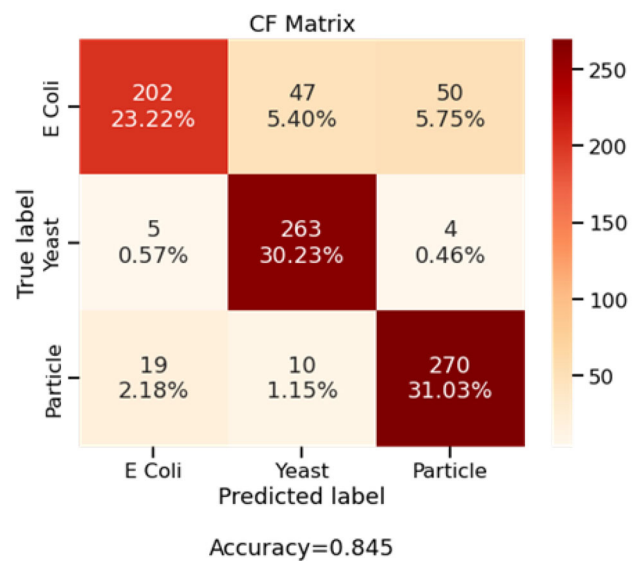


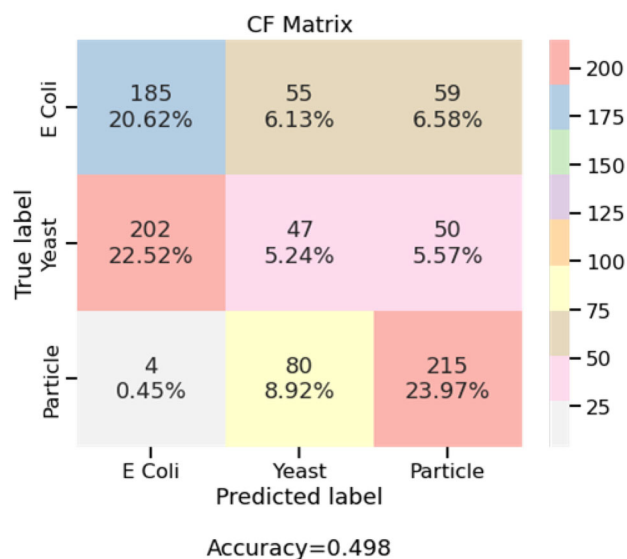
Fig. 9 Confusion matrix of the proposed model representing the classification of E. Coli, yeast, and particles

recognition of the Bacteria from water using the fine-tuned model with variable parameters. The trained model is the best fit and provides accurate results on microscopic images due to Transfer Learning involvement.

Our proposed research model mainly focuses on the detection of Escherichia Coli and Yeast while any other material found in the water is a particle and makes them detection as early as possible in terms of accuracy and efficiency. Detecting the Bacteria categorizes enables deep learning to provide an effective platform that can be transformative for a variety

Table 3 Performance evaluation of proposed model in comparison to the state-of-the-art

Deep Features	Accuracy	Kappa
YoloV5	0.331	0.303
SSD-Net	0.498	0.247
M-YoloV5	0.846	0.768

**Fig. 10** Confusion matrix of SSD-Net representing the classification of E. Coli, yeast, and particles

of applications in microbiology by reducing the detection time and automated counting of the colonies without labeling or the need for an expert. In the future, the research can be extended by exploring multi-class detection to identify a broader range of bacterial pathogens in water samples. Optimizing the model for real-time deployment on edge devices would enable on-site water quality monitoring. Additionally, the research can explore multi-modal analysis by incorporating other imaging modalities. The techniques developed in this research can be explored in other fields beyond water quality assessment, and efforts can be made to quantify the uncertainty in detection results for more reliable decision-making.

Future work

In the future, the research can be extended by exploring multi-class detection to identify a broader range of bacterial pathogens in water samples. Optimizing the model for real-time deployment on edge devices would enable on-site water quality monitoring. Transfer learning can be investigated to enhance the model's performance with limited labeled data. Developing a comprehensive water quality

monitoring system integrating the proposed model with sensor networks would allow continuous monitoring and timely alerts for contamination events. Additionally, the research can explore multi-modal analysis by incorporating other imaging modalities. Creating a user-friendly mobile application would empower users to capture water sample images and receive contamination risk scores. The techniques developed in this research can be explored in other fields beyond water quality assessment, and efforts can be made to quantify the uncertainty in detection results for more reliable decision-making.

Acknowledgements The authors would like to thank HEC- National Research Program for Universities of Pakistan (*ProjectNumber* : 7794).

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Aghalari, Z., Dahms, H.U., Sillanpää, M., et al.: Effectiveness of wastewater treatment systems in removing microbial agents: a systematic review. *Glob. Health* **16**, 1–11 (2020)
2. Bridge, J.W., Oliver, D.M., Chadwick, D., et al.: Engaging with the water sector for public health benefits: waterborne pathogens and diseases in developed countries. *Bull. World Health Organ.* **88**, 873–875 (2010)
3. Chandnani, G., Gandhi, P., Kanpariya, D., et al.: A comprehensive analysis of contaminated groundwater: Special emphasis on nature-ecosystem and socio-economic impacts. *Groundw. Sustain. Dev.* **100813**, (2022)
4. Chayadevi, M., Raju, G.: Extraction of bacterial clusters from digital microscopic images through statistical and neural network approaches. In: *Proceedings of International Conference on Advances in Computing*, Springer, pp 1091–1099 (2012)
5. Cheema, M.N., Nazir, A., Sheng, B., et al.: Liver extraction using residual convolution neural networks from low-dose ct images. *IEEE Trans. Biomed. Eng.* **66**(9), 2641–2650 (2019)

6. Deng, L., Zhong, Y., Wang, M., et al.: Scale-adaptive deep model for bacterial raman spectra identification. *IEEE J. Biomed. Health Inform.* **26**(1), 369–378 (2021)
7. Engineers, C.: Long term 2 enhanced surface water treatment rule toolbox guidance manual. US Environmental Protection Agency, Office of Water (2010)
8. Galar, A., Yuste, J., Espinosa, M., et al.: Clinical and economic impact of rapid reporting of bacterial identification and antimicrobial susceptibility results of the most frequently processed specimen types. *Eur. J. Clin. Microbiol. Infect. Dis.* **31**, 2445–2452 (2012)
9. Gopinath, S.C., Tang, T.H., Chen, Y., et al.: Bacterial detection: From microscope to smartphone. *Biosens. Bioelectron.* **60**, 332–342 (2014)
10. Hosang, J., Benenson, R., Schiele, B.: Learning non-maximum suppression. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4507–4515 (2017)
11. Khayyat, M.M.: Improved bacterial foraging optimization with deep learning based anomaly detection in smart cities. *Alex. Eng. J.* **75**, 407–417 (2023)
12. KhoKhar, F.A., Shah, J.H., Khan, M.A., et al.: A review on federated learning towards image processing. *Comput. Electr. Eng.* **99**, 107818 (2022)
13. Marston, H.D., Dixon, D.M., Knisely, J.M., et al.: Antimicrobial resistance. *JAMA* **316**(11), 1193–1204 (2016)
14. Men, H., Wu, Y., Gao, Y., et al.: Application of support vector machine to heterotrophic bacteria colony recognition. In: 2008 International conference on computer science and software engineering, IEEE, pp 830–833 (2008)
15. Mortier, T., Wieme, A.D., Vandamme, P., et al.: Bacterial species identification using maldi-tof mass spectrometry and machine learning techniques: a large-scale benchmarking study. *Comput. Struct. Biotechnol. J.* **19**, 6157–6168 (2021)
16. Mosleh, M.A., Manssor, H., Malek, S., et al.: A preliminary study on automated freshwater algae recognition and classification system. In: *BMC bioinformatics*, BioMed Central, pp 1–13 (2012)
17. Nazir, A., Cheema, M.N., Sheng, B., et al.: Off-enet: An optimally fused fully end-to-end network for automatic dense volumetric 3d intracranial blood vessels segmentation. *IEEE Trans. Image Process.* **29**, 7192–7202 (2020)
18. Nazir, A., Cheema, M.N., Sheng, B., et al.: Spst-cnn: Spatial pyramid based searching and tagging of liver's intraoperative live views via cnn for minimal invasive surgery. *J. Biomed. Inform.* **106**, 103430 (2020)
19. Nnachi, R.C., Sui, N., Ke, B., et al.: Biosensors for rapid detection of bacterial pathogens in water, food and environment. *Environ. Int.* **166**, 107357 (2022)
20. Noble, R.T., Weisberg, S.B.: A review of technologies for rapid detection of bacteria in recreational waters. *J. Water Health* **3**(4), 381–392 (2005)
21. Organization, W.H.: The world health report 2002: reducing risks, promoting healthy life. World Health Organization (2002)
22. Organization, W.H., et al.: Water, sanitation, hygiene and health: a primer for health professionals. Tech. rep., World Health Organization (2019)
23. Osman, M.K., Ahmad, F., Saad, Z., et al.: A genetic algorithm-neural network approach for mycobacterium tuberculosis detection in ziehl-neelsen stained tissue slide images. In: 2010 10th international conference on intelligent systems design and applications, IEEE, pp 1229–1234 (2010)
24. Papa, F., Crétau, J.F., Grippa, M., et al.: Water resources in africa under global change: monitoring surface waters from space. *Surv. Geophys.* **44**(1), 43–93 (2023)
25. Ray, S.S., Chen, S.S., Li, C.W., et al.: A comprehensive review: Electrospinning technique for fabrication and surface modification of membranes for water treatment application. *RSC Adv.* **6**(88), 85495–85514 (2016)
26. Rhoads, D.D., Wolcott, R.D., Sun, Y., et al.: Comparison of culture and molecular identification of bacteria in chronic wounds. *Int. J. Mol. Sci.* **13**(3), 2535–2550 (2012)
27. Sandgren, A., Strong, M., Muthukrishnan, P., et al.: Tuberculosis drug resistance mutation database. *PLoS Med.* **6**(2), e1000002 (2009)
28. Tamiev, D., Furman, P.E., Reuel, N.F.: Automated classification of bacterial cell sub-populations with convolutional neural networks. *PLoS ONE* **15**(10), e0241200 (2020)
29. Trattner, S., Greenspan, H., Tepper, G., et al.: Automatic identification of bacterial types using statistical imaging methods. *IEEE Trans. Med. Imaging* **23**(7), 807–820 (2004)
30. Visitsattapongse, S., Bunkum, M., Pintavirooj, C., Paing, M. P.: A deep learning model for bacterial classification using Big Transfer (BIT). *IEEE Access*, IEEE (2024). <https://doi.org/10.1109/ACCESS.2024.3358671>
31. Centers for Disease Control and Prevention (CDC) and Water, Healthy and others: Current waterborne disease burden data & gaps (2017)
32. Xiaojuan, L., Cunshe, C.: An improved bp neural network for wastewater bacteria recognition based on microscopic image analysis. *WSEAS Trans. Comput.* **8**(2), 237–247 (2009)
33. Xu, L., Xie, J., Jiang, T., et al.: Red tide algae classification using svm-snp and semi-supervised fcm. In: 2010 2nd International Conference on Education Technology and Computer, IEEE, pp V1–389 (2010)
34. Yasmine, G., Maha, G., Hicham, M.: Overview of single-stage object detection models: from yolov1 to yolov7. In: 2023 International Wireless Communications and Mobile Computing (IWCMC), IEEE, pp 1579–1584 (2023)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Fahad Ahmed Khokhar received the B.S. degree in Software Engineering, in 2018, and the M.S. degree in Computer Science, in 2016, from the COMSATS University, Islamabad, Pakistan. He is currently pursuing the Ph.D. degree with the University of Florence, Italy. From 2020 to 2022, he worked as a Research Assistant with the Department of Computer Science, COMSATS University, Islamabad, Pakistan. His research interests include anomaly detection, security and safety, computer vision, and implement appropriate architectures or software in the domain of critical systems.



Jamal Hussain Shah received the Ph.D. degree in pattern recognition from the University of Science and Technology China, Hefei, China. Since 2008, he has been in the education field. He is currently serving as Tenured Associate Professor at COMSATS University Islamabad, Wah Campus, Pakistan. His areas of specialization are automation and pattern recognition. He has over 60 publications in IF, SCI and ISI journals and in national and international conferences. His research interests

include deep learning, algorithms design and analysis, machine learning, image processing, and big data.



Rabia Saleem is an Assistant Professor in the Department of Information Technology, Government College University, Faisalabad, Pakistan. Renowned for her expertise in computer science and machine learning. She earned her Bachelor's degree in Computer Science (BSCS) from a prestigious institution, followed by a Master's degree (MSCS) specializing in advanced algorithms. Motivated by a passion for cutting-edge research, she pursued a Ph.D. in Computer Science, focusing her

doctoral work on machine learning applications. Her research contributions include novel algorithms for segmentation techniques and pattern recognition, advancing the field's understanding of artificial intelligence and machine learning. As an Assistant Professor, she combines her academic pursuits with teaching, mentoring students, and guiding research projects that address real-world challenges in technology and AI.



Anum Masood received her Ph.D. degree in Computer Science and Engineering from Shanghai Jiao Tong University, Shanghai, China in 2019. Her B.Sc. and M.Sc. degrees in Computer Science are from the COMSATS University Islamabad, Islamabad, Pakistan. Currently, she is a research fellow at Harvard Medical School and a postdoctoral fellow affiliated with the Norwegian University of Science and Technology, and PET Centre, St. Olav's Hospital, Norway. From 2014 to 2020, she was

a Lecturer at the Department of Computer Science at COMSATS University Islamabad. She also worked as a Visiting Researcher with the Institute of Neuroscience and Medicine (INM-2), Forschungszentrum Jülich; Institute for Cardiogenetics, University of Luebeck, Germany and Liverpool John Moores University, U.K. Her research interests include medical image analysis, automated cancer detection, machine learning, and image processing.