SIN-CHN
SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 9(1), 609.
https://doi.org/10.62617/mcb609

Article

Rapid detection of food microorganisms from the perspective of cellular and molecular biomechanics leveraging biotechnology and computer vision

Yan Huang

Food Department, Minbei Vocational and Technical College, Nanping 353000, China; mbu005@163.com

CITATION

Huang Y. Rapid detection of food microorganisms from the perspective of cellular and molecular biomechanics leveraging biotechnology and computer vision. Molecular & Cellular Biomechanics. 2025; 9(1): 609. https://doi.org/10.62617/mcb609

ARTICLE INFO

Received: 24 October 2024
Accepted: 11 November 2024
Available online: 10 January 2025

Abstract: Food microbiological detection is an important part of food safety management. Understanding the behavior and characteristics of microorganisms at the cellular and molecular level can enhance the detection process. To quickly and effectively detect food microorganisms, a rapid microorganisms detection method based on biotechnology and computer vision is proposed. Firstly, food bacterial strains are cultivated based on biotechnology and sample data is prepared. At the cellular level, this involves understanding the growth kinetics and metabolic processes of the microorganisms. Secondly, a microorganisms classification detection model is proposed based on residual neural networks, and transfer learning and attention mechanisms are introduced to optimize the model. By mimicking the way cells and molecules interact and signal, these techniques can help the model better recognize and classify different microbial species. Considering the problem of insufficient detection of large-scale complex scenes, an improved object detection model is proposed, which introduces a lightweight model to replace the backbone feature network and uses deep separation convolution to replace ordinary convolution, thereby raising the training accuracy of the model. In the classification model experiment, the research model shows better model loss performance and microorganisms detection accuracy in both low-density and high-density microbial scenarios. In the analysis of object detection models, compared to other models, the research model has smaller losses. In large-scale scenes and multi-feature large-scale scenes, the research model has losses of 0.12 and 3.56, respectively, which are better than other models. In addition, in common microorganisms detection, the high accuracy of 99.75% for detecting Escherichia coli indicates the model's proficiency in recognizing the specific cellular and molecular characteristics of this microorganism, providing significant technical references for ensuring food safety and efficient microorganism detection from a cellular and molecular biomechanics perspective.

Keywords: biotechnology; computer vision; residual neural network; detection model; attention mechanism; cellular and molecular level

1. Introduction

Food safety issues are a hot and key concern for people, and food safety issues caused by microbial contamination have received widespread attention from society. At present, common harmful microorganisms in food include Escherichia coli, Salmonella, mold, etc. Harmful microorganisms can lead to the risk of food poisoning, death, and even cancer in humans [1]. Therefore, in recent years, strengthening microbial detection in food has been the key to ensuring food safety. Currently, there are many methods for microorganisms detection, including microbial culture method, biological color labeling method, etc. [2]. These aspects have problems such as low detection efficiency, small detection range, and high detection cost in practical applications. In recent years, computer vision technology

has made rapid progress, and through more advanced image recognition systems, it can quickly identify and determine the type of target, which has become an important development point in food microbiology detection technology [3]. To solve the problems of low efficiency and poor detection effect of traditional microbial detection technology, a rapid detection method of food microorganisms based on biotechnology and computer vision technology was proposed. There are two innovations in the research technology. One is based on biotechnology, which collects and makes microbial experimental samples, and combines computer vision technology for data training to ensure the reliability of the technology. At the same time, the transfer learning strategy, target detection algorithm and other technologies are introduced to enhance the image feature analysis, so as to further improve the effect of the technology in microbial recognition. This study provides an important technical reference for food safety management and efficient microbial detection.

The research is structured into four parts. The part 1 illustrates common microorganisms detection technologies and computer vision technologies, and discusses their applications in the field of food safety. The part 2 is to analyze food microorganisms, prepare microbial samples, and construct detection models based on the scale of microorganisms. The part 3 is to apply the mentioned technology to specific scenarios and verify the practical application effects of the two microorganisms detection models in food detection. The part 4 summarizes and analyzes the entire text, and elaborates on the improvement direction of the research.

2. Related work

In recent years, food health issues have attracted people's attention, and more and more scholars are conducting research on food microbiological testing. Xing et al. [4] conducted research on food detection and introduced typical methods for pathogen capture, isolation, and detection to improve detection efficiency. Then, the key applications of rapid microorganisms detection based on microfluidic biosensors were introduced in detail, achieving the capture, recognition, and counting of bacteria. Comparing this technology with traditional food microbiology detection techniques, the research technology has higher sensitivity and detection effectiveness, and is suitable for more food safety detection scenarios. Xiao et al. [5] studied a color fluorescence synthesis detection technique to improve the detection efficiency of microorganisms. This technology could detect microorganisms in the food environment through probes, effectively identifying the degree of food spoilage and safety effects. Finally, this technology was compared with traditional food microbiology detection techniques, and compared to other technologies, the accuracy of the studied technology was higher and the detection effect was more outstanding. Guo et al. [6] found that with the development of the globalized economy, food pollution has become increasingly severe. To quickly and effectively detect food safety, research is being conducted on the detection of food microorganisms based on traditional microbiological techniques. A novel microorganisms detection technique using surface enhanced Raman spectroscopy was adopted and tested in different foods. The results showed that this technology could effectively detect food and aquatic organisms, and had good application prospects.

With the continual advancement of computer vision technology, advanced and efficient intelligent microorganisms detection technology has been broadly utilized in the field of food safety. Yang et al. [7] found that traditional food microorganisms testing techniques cannot detect multiple types of food. A highly efficient microorganisms detection technology based on computer vision was proposed in this study. This technology utilized a pathogen identification system using a paper color array to identify and analyze pathogens in food. The color change numbers between different pathogens were input into a neural network system, and deep learning was used to identify and test different food microorganisms. Through model training, compared to traditional microorganisms detection techniques, the overall effectiveness of the research technology was more outstanding. Zhang et al. [8] raised a deep learning detection model with computer vision, which can quickly detect microorganisms through microscopic imaging systems by training various microorganisms, bacteria, parasites, and fungi. Compared with traditional microorganisms techniques, research techniques had better overall effectiveness and higher detection efficiency. Firouz et al. [9] found that traditional microorganisms detection techniques had low efficiency and limited adaptability to various scenarios, which cannot meet the requirements of agricultural food safety testing. So, combining computer vision technology, a deep learning food safety detection technology was proposed. This technology analyzed and detected food safety by detecting defects in food samples and classifying image data. Finally, the technology was applied to specific agricultural product safety testing processes, and it had excellent detection results. Finally, Martín et al. [10] proposed a novel method for counting bacteria and yeast in microbial biological products using digital image processing. A database method in Python language that utilizes basic digital image processing operations such as contour detection, morphological operations, and statistical analysis was proposed to achieve detection and evaluation of microorganisms. The application of this technology in food safety processes showed excellent performance.

In summary, microorganisms detection plays a crucial role in food safety. At present, microorganisms detection technologies in food mainly include molecular microbial learning method, traditional technology method, etc., all of which have certain limitations and average detection results. In recent years, the usage of computer vision in the microorganisms detection has significantly promoted the development of food safety. Based on biotechnology and computer vision, an efficient food microorganisms detection method is proposed to meet the requirements of food safety detection.

3. Materials and methods

This section mainly analyzes the microorganisms in food and produces microbial image data. At the same time, classification detection models and object detection models are constructed based on the scale of microorganisms for microorganisms detection in different scenarios.

3.1. Materials

Experimental equipment: Olympus CX31 scientific microscope, provided by Tokyo Olympus, Japan; Multi specification culture dishes produced by Zhejiang Runlan Technology Co., Ltd. in Taizhou, Zhejiang, China; Multi specification test tubes produced by Zhejiang Runlan Technology Co., Ltd., located in Taizhou, Zhejiang, China; Mettler ME balance, provided by Mettler Toledo, Shanghai, China; Multi specification measuring cups and culture dishes produced by Zhejiang Runlan Technology Co., Ltd. in Taizhou, Zhejiang, China; HEPA type sterile operating table, provided by Suzhou Bolanke Instrument Equipment Co., Ltd., Suzhou, Jiangsu, China; Thermo Scientific371 insulated box, provided by Thermo Scientific, Waltham, Massachusetts, United States; German GFL type constant temperature shaker provided by German GFL company, Borgweiler, Germany; Thermostatic box, provided by Grenier Precision, China, Jiangsu, Suzhou.

Experimental materials: 400 grams of fresh pork, 250 grams of chicken, duck, and beef each, all provided by Sichuan New Hope Food Co., Ltd., Chengdu, Sichuan, China.

Experimental reagent: sterile physiological saline, provided by Shanghai Yueteng Biotechnology Research Institute, Shanghai, China; Agar agar plate culture medium, provided by Shanghai Nordic Biotechnology Co., Ltd., Shanghai, China; Methylene blue indicator, provided by Sichuan Weiqi Biotechnology Co., Ltd., Chengdu, Sichuan, China; Experimental purified water, provided by Hangzhou Wahaha Beverage Co., Ltd. in Hangzhou, Zhejiang, China; Polyformaldehyde, provided by Shandong Aldehydes Chemical Co., Ltd., Linyi, Shandong, China; Ionized water, provided by Gaide Chemical, China, Jiangsu, Suzhou.

3.2. Method

3.2.1. Sample preparation based on biotechnology

Microorganisms detection technology plays a crucial role in areas such as healthcare, food safety, wastewater treatment, and biological research, and is one of the most direct and effective methods for evaluating food safety. In recent years, food safety has attracted much attention. Traditional microorganisms detection techniques are cumbersome, time-consuming, and have average detection results, which can no longer meet food safety requirements [11]. A smart food microbiology detection technology based on computer vision technology is proposed. This technology utilizes traditional biotechnology for microbial cultivation to produce food microbiological image data. Then, image data is recognized through computer vision models, and rapid detection of microorganisms is achieved through object classification and recognition [12]. The technical framework of the entire technology is shown in **Figure 1**.

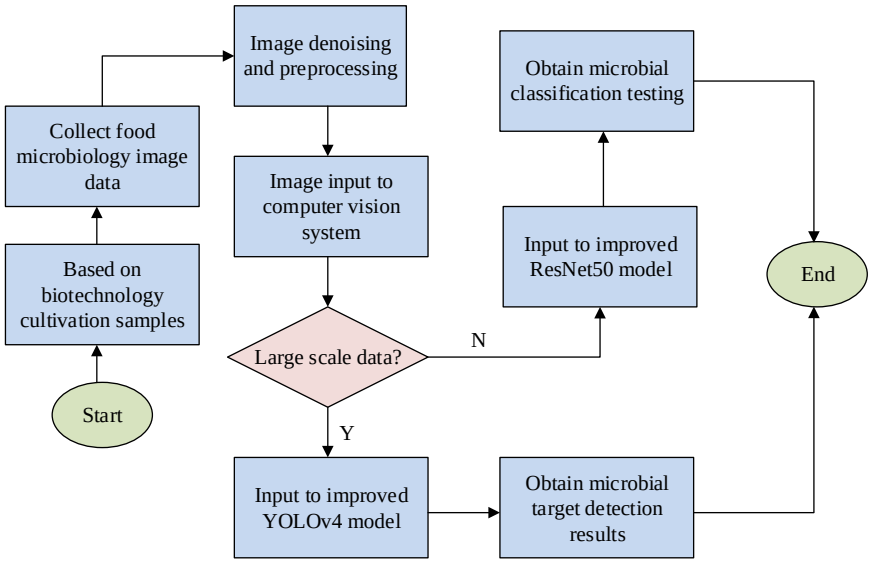

Figure 1. Framework of food microorganisms detection technology.

At present, common microorganisms in food include Escherichia coli, Salmonella, Staphylococcus aureus, etc. To effectively detect food microorganisms, microbial sample data will be produced based on microbial technology for subsequent model training. The main tools used in data production include incubators, which are German Thermo Scientific models, German GFL model constant temperature shaker, Olympus CX31 microscope image acquisition system, multi specification culture dishes and test tubes, Mettler ME scale and multi specification measuring cup and HEPA model sterile operating table.

To establish an experimental training sample database, four major categories of food are selected as food research objects: meat, fresh fruits and vegetables, grains and flour, and drinking water. Firstly, four categories of food that need to be tested will be prepared, and all food samples will be provided by Sichuan New Hope Food Company [13]. The national standard method is used to test the four major categories of food, and the entire data production process is shown in **Figure 2**.

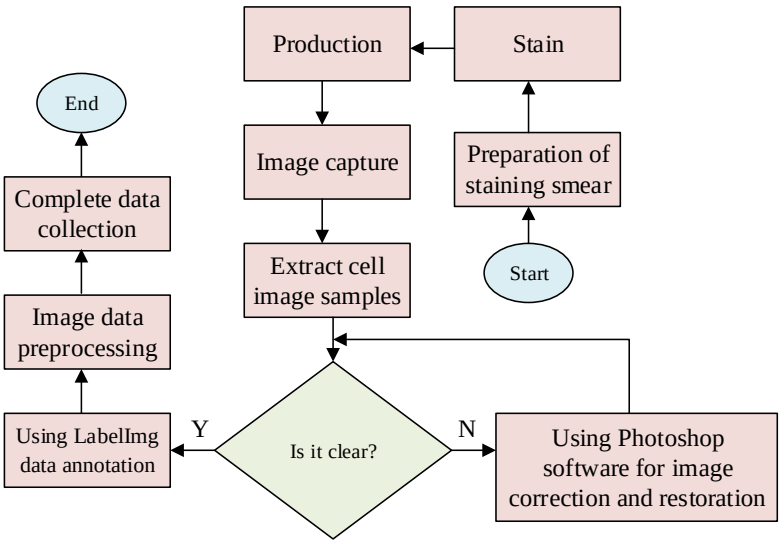

Figure 2. Sample data production process.

For producing training samples for bacterial cells, the following steps can be followed. Firstly, the preparation of the smear involves weighing 25 g of meat as a sample and placing it in 225 mL of sterile physiological saline. The sample is then prepared at 8000 r/min to form a 1:10 sample homogenization. If the sample is ordinary pure water during production, it is necessary to use a sterile straw to take 25 mL of the sample and put it into 225 mL of sterile physiological saline. It stirs to make a 1:10 sample homogenization. Then it estimates the contamination of the sample, takes 1mL of the sample and places it in a sterile petri dish, which needs to be cooled to 46 ℃ for uniform shaking. Until the agar solidifies, it places the plate in a constant temperature incubator at 36 ℃ for cultivation, and forms colonies. Finally, it inoculates onto sterile plates for refrigeration.

Finally, it is necessary to stain the microorganisms sample. Considering its high moisture content, the living dye methylene blue is used for staining. The stained smear is placed on the microscope slide and it is covered with a glass cover to fix the bacteria on the slide. Finally, it takes pictures and places the prepared slide under the lens of a microscope, using image capture equipment to capture images of bacterial cells. Finally, the collected microbial characteristic images will be processed for clarity and data preprocessing, and the final experimental sample data will be obtained.

3.2.2. Construction of microbial classification model based on computer vision

To effectively detect microorganisms in food, a computer vision-based microbial classification detection technology is proposed. Compared to manual methods, computer vision is more efficient in microbial detection and has higher recognition accuracy through extensive training in microbial image recognition. In the study, an improved Residual Neural Network (ResNet50) model was used as a visual recognition technique, which has the advantage of effectively solving the degradation problem of deep neural networks. It can extract richer image features, with suitable depth for relatively efficient training, high accuracy in various visual recognition tasks such as classification and detection, and wide applicability [14]. In addition, the study introduced transfer learning strategies to improve the ResNet50 model parameters and enhance the model's ability to extract edge contour details from images. In the ResNet50 model, it is assumed that the input image being studied is X_o, which is a three-dimensional tensor of $H \times W \times C$, where H represents the height of the image, W refers to the width of the image, and C denotes the number of channels in the image. The input expression of the image is shown in Equation (1).

$$X_o = H \times W \times C \tag{1}$$

The study first extracts image features through convolutional layers. It assumes the convolution kernel size used in the study is $K \times K$ and the step size is S. By using convolution operation, the feature map F_1 is obtained, and its dimension K_d is shown in Equation (2).

$$K_d = (\frac{H}{S}) \times (\frac{W}{S}) \times D_1 \tag{2}$$

In Equation (2), D_1 denotes the channels in the feature map. In the model, each residual block consists of two paths, the backbone path and the skip connection path. Assuming there are L residual blocks in the study, the input of the l residual block is F_{l-1}, and the output is F_l. The backbone path consists of two consecutive convolutional layers and a batch normalization layer, where each convolutional layer has a 1×1 convolution kernel and a 3×3 convolution kernel [15]. The relationship between the output and input of each residual block is shown in Equation (3).

$$F_l = F_{l-1} + H_l(F_{l-1}) \tag{3}$$

In Equation (3), H_l represents the mapping function of the residual block. Finally, the study connects the output of the global average pooling layer to a fully connected layer that maps feature vectors to the probability distribution of microbial categories. Assuming the weight matrix of the fully connected layer is W_i and the bias vector is b, then the output of the fully connected layer is shown in Equation (4).

$$Y_o = W_i X_o + b \tag{4}$$

In Equation (4), Y_o is a vector of $1 \times K$, representing the probability distribution of microbial categories. Considering that the ResNet50 model needs to handle texture and edge details during training, which increases the computational complexity of the model, a Transfer Learning (TL) strategy is introduced for parameter optimization to improve model efficiency [16]. Assuming the study has a pre trained model M_1, the study can use the parameters of M_1 as the initial parameters of the ResNet50 model. For the previous convolutional layers and residual blocks, research can fix their parameters unchanged and only train the parameters of the final fully connected layer. By using a dataset to fine tune the model, it can better adapt to microbial classification tasks [17].

In addition, considering the lack of attention to important features in the ResNet50 model, a visual attention mechanism (VAM) is introduced for optimization. Assuming that the input image of the study is segmented to obtain U_R regions, each region's feature is represented as X_i, as shown in Equation (5).

$$U_R = \{X_1, X_2, X_3, \ldots, X_i\} \tag{5}$$

In Equation (5), $i = 1,2,\ldots,R$. The study uses a mapping function A to calculate attention weights for each region, namely $A(X_i)$. The mapping function can be a convolution operation or a fully connected layer. To obtain a normalized attention weight vector W_i, the study can use the $softmax$ function to normalize the mapped weights, and the weight vector expression is shown in Equation (6).

$$W_i = \text{softmax}(A(X_1), A(X_2), \ldots, A(X_R)) \tag{6}$$

The attention weight vector W_i obtained from the study will be multiplied by the feature vectors of each region to obtain the weighted feature representation, as shown in Equation (7).

$$X_i^{'} = W_i \odot X_i \tag{7}$$

In Equation (7), $\odot$ represents element wise multiplication operation. The research will fuse the weighted feature representations, and the final feature

representation can be obtained through summation or concatenation operations, as shown in Equation (8).

$$X' = \sum_{i=1}^{R} X_i' \tag{8}$$

By constructing the above mathematical model, research can achieve effective classification and detection of food microorganisms. The construction process of microbial classification model is shown in **Figure 3**.

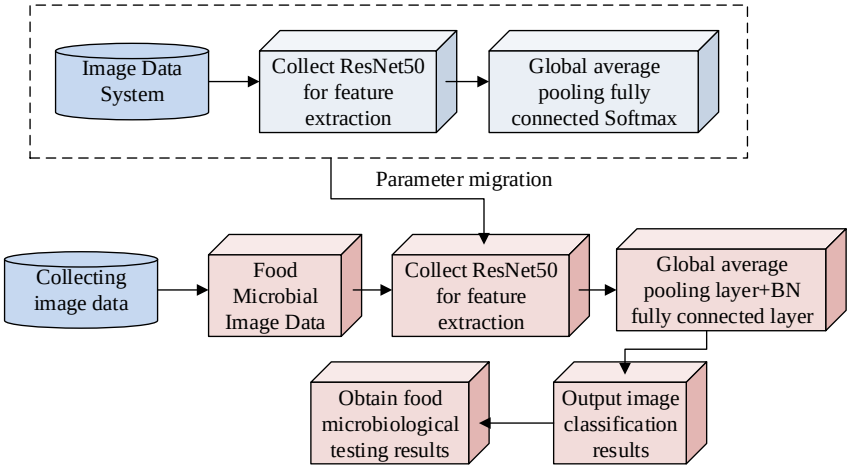

Figure 3. Construction process of microbial classification model.

3.3. Construction of microbial object detection model based on improved computer vision

In food microbiology detection, the microbial classification detection model is suitable for microbiology detection in small field environments. For complex scenarios with large microbial populations, it is not possible to use microbial classification models for effective detection. Therefore, based on the fourth generation object detection system You Only Look Once version 4 (YOLOv4), an object detection model suitable for complex and large-scale microbial scenes is proposed [18]. It assumes the input image for the study is $X \in \mathbb{R}^{H_r \times W_r \times C_r}$, where H_r represents the height of the image, W_r represents the width of the image, and C_r represents the number of channels in the image. In YOLOv4, a backbone network is used to extract image features. Backbone networks typically use operations such as convolutional layers and pooling layers to gradually extract advanced features from images. The research assumes that the output of the backbone network is shown in Equation (9).

$$f \in \mathbb{R}^{H' \times W' \times C'} \tag{9}$$

In Equation (9), H' denotes the height of the feature map, W' denotes the width of the feature map, and C' denotes the amounts of channels in the feature map. In YOLOv4, a Neck part is introduced to further adjust the dimensionality of feature maps to adapt to object detection tasks. The output of the research hypothesis Neck is shown in Equation (10).

$$f' \in \mathbb{R}^{H'' \times W'' \times C''} \tag{10}$$

In Equation (10), H'' denotes the height of the adjusted feature map, W'' denotes the width of the adjusted feature map, and C'' denotes the amounts of channels in the adjusted feature map. In YOLOv4, a head section is used to complete the object detection task. The research assumes that the output of the head part is shown in Equation (11).

$$Y \in \mathbb{R}^{S \times S \times (B \times 5 + C')} \tag{11}$$

In Equation (11), S represents the size of the output detection map, B represents the amounts of bounding boxes detected by each detection box, and C' represents the amounts of categories. In the head section of YOLOv4, research is conducted on using convolutional layers to detect the position and category of targets. For each bounding box, four coordinates (x, y, w, h) are used to represent its position, where (x, y) means the center coordinate of the bounding box and (w, h) means the width and height of the bounding box [19]. Considering the complexity of microbial feature extraction, MobileNet is used to replace the CSPDarknet 53 network for feature extraction in the study. At the same time, depthwise separable convolution is used to replace the 3x3 ordinary convolution of the model, thereby reducing the training parameters of the model [20]. The principle of depthwise separable convolution is to perform convolution operations on each input channel separately, which can effectively extract features within the channel. Its impact on the model includes improved computational efficiency: the number of parameters and computational complexity of depthwise separable convolutions are significantly reduced compared to ordinary 3 × 3 convolutions [21]. Secondly, it is important to prevent overfitting: reducing parameters can also help alleviate overfitting issues and improve the model's recognition accuracy on new data. The principle of depthwise separable convolution is indicated in **Figure 4**.

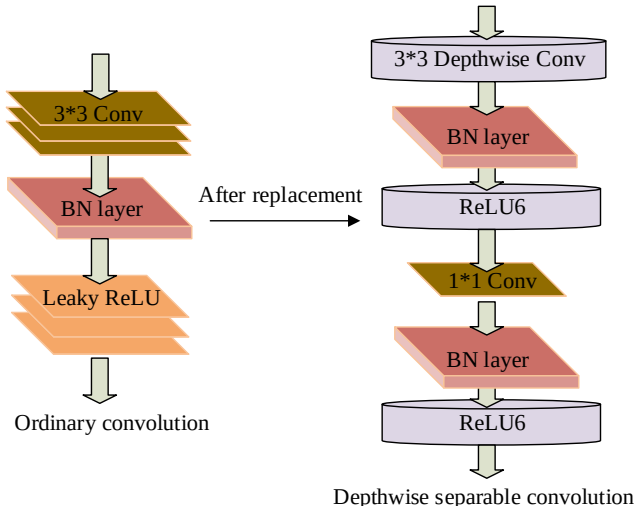

Figure 4. Depthwise separable convolution.

In **Figure 4**, ordinary convolution has 9 times the computational complexity of depthwise separable convolution. A separable convolution kernel size of 3 × 3 is

used to cut down the computational complexity, and the RELU6 activation function is used to optimize the output parameters to ensure image output accuracy [22]. The width and height of the feature matrix is D_G, the amount of input channels is M, the size of the convolution kernel is D_K, the amount of output channels is N, and the width and height of the output feature matrix is D_F. Therefore, the ordinary convolution operation is shown in Equation (12) [23].

$$C_c = D_K \times D_K \times N \times D_F \times D_F \tag{12}$$

Depthwise separable convolution is introduced to replace traditional convolution, and its computational complexity is shown in Equation (13).

$$C_d = D_K \times D_K \times M \times D_F \times D_F \tag{13}$$

In deep convolution operations, each feature map channel is relatively independent and requires pointwise convolution calculation, as shown in Equation (14) [24].

$$C_p = M \times N \times D_F \times D_F \tag{14}$$

The total computational cost of the final depthwise separable convolution is shown in Equation (15).

$$C_{ds} = (D_K \times D_K + N) \times M \times D_F \times D_F \tag{15}$$

In addition, in complex and large-scale microbial identification, to enhance the model's ability to extract microbial features, attention mechanisms are introduced to promote the model's attention to microbial features. The attention mechanism is added as shown in **Figure 5**.

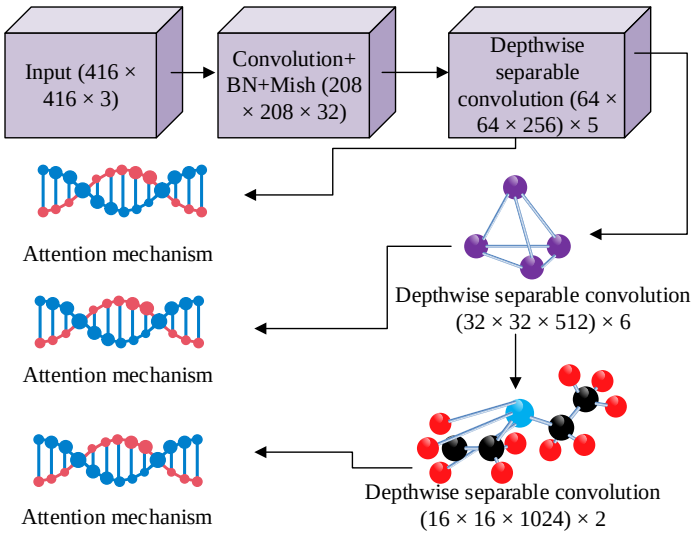

Figure 5. Schematic diagram of adding attention.

Through the above optimization, the construction of the object detection model has been completed. This model adopts YOLOv4 as the foundation, introduces MobileNet as the backbone feature extraction network, and uses attention mechanism to promote the model's ability to extract microbial features, ensuring better feature extraction ability and improving detection performance.

4. Experimental analysis of food microbiology detection

This section will analyze the model performance and actual detection results of the two proposed food microorganism detection models. The main evaluation indicators include precision, recall rate, detection accuracy, loss value, and other indicators.

4.1. Experimental analysis of microbial classification model

To assess the effectiveness of the raised classification model, self-made food microorganism samples were selected as experimental sample data, including 25 types of microorganisms, 15632 image data, and a ratio of 7:3 between the training and testing sets. Experimental analysis was conducted on the Pytorch platform. The initial parameters of the improved ResNet50 model are indicated in **Table 1**.

Table 1. Model initial parameters.

Parameter indicator type	Numerical value
Requires_grad	true
Learning rate	1e-3
EPOCHS	30
BATCH_SIZE	16

CNN and Gradient Boosting Decision Tree (GBDT) optimized ResNet50 were introduced as the testing benchmark. The Precision Recall (P-R) curve, detection accuracy, and loss value were introduced as evaluation indicators. The larger the area under the P-R curve, the better the model detection. The comparison of loss performance among different models is shown in **Figure 6**.

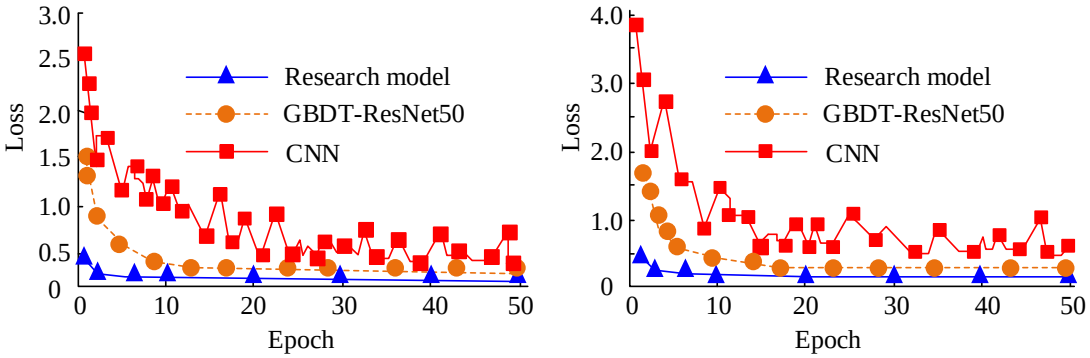

Figure 6. Performance loss comparison under different models.

Figure 6a showcases the model loss results under low-density microorganisms. Among the three models, the traditional CNN model had the highest loss and showed significant fluctuations in 50 iterations. Overall, the research model performed the best and converged the fastest. After 50 iterations, the loss of the research model, CBDT ResNet50, and CNN were 0.48, 0.24, and 0.08, respectively. **Figure 6b** shows the model loss results under high-density microorganisms. The results still showed the fastest convergence of the research model and the lowest loss. The loss

value of the research model was 0.09 after 50 iterations. **Figure 7** shows the P-R curves of different models.

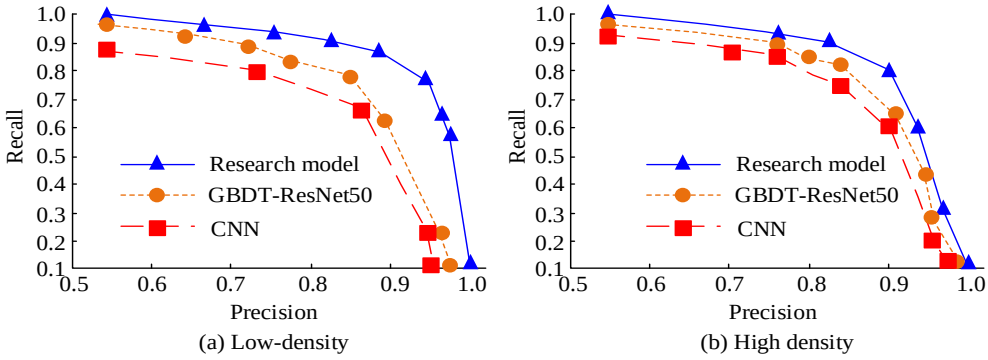

Figure 7. Comparison of P-R curves for different models.

Figure 7a,b show the detection results of low-density and high-density microorganisms, respectively. At low density, the CBDT ResNet50 model, which has the largest surrounding area and performs second, had better model detection performance. Compared to CBDT ResNet50 and CNN, the research model showed a detection performance improvement of 8.65% and 14.65%. At high density, the overall performance of the research model was still the best, with a detection performance improvement of 6.65% and 10.35% compared to CBDT ResNet50 and CNN. **Figure 8** shows the microorganisms detection results under different models.

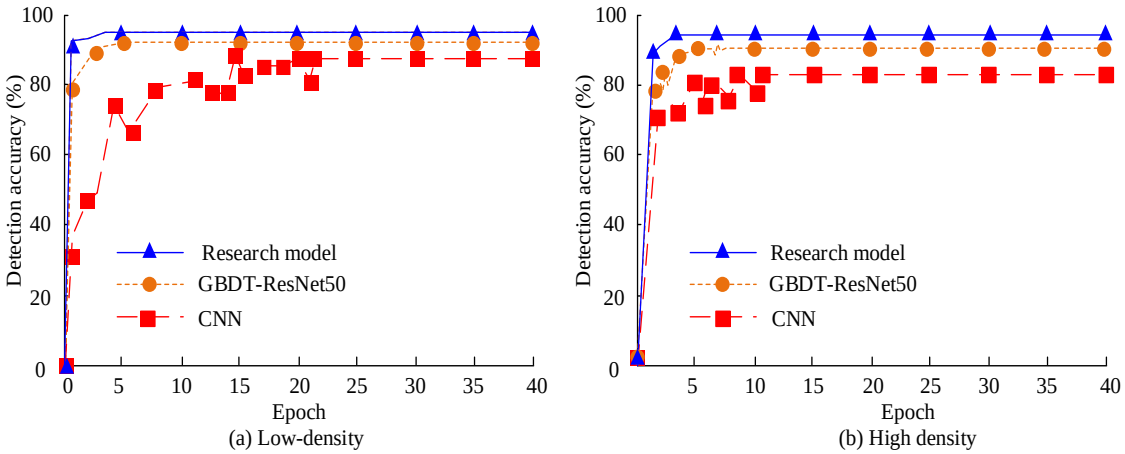

Figure 8. Comparison of detection accuracy under different models.

Figure 8a,b show the detection accuracy under low-density and high-density microorganisms, respectively. According to the test results, the accuracy of microorganisms detection decreased in high-density compared to low-density, but the best performing model was the research model. Meanwhile, traditional CNN had the worst detection performance, with significant detection fluctuations in both low-density and high-density detection. At high and low densities, the detection accuracy of the research model was 98.65% and 97.25%, respectively, which was better than the other two models.

4.2. Experimental analysis of microorganisms object detection model

In large-scale and complex food microorganisms detection, the proposed object detection model was used for microorganisms testing experiments. The Batch size value was set to 48 and the amounts of iterations to 100. At the same time, YOLOv4 was introduced, which replaced the backbone feature network MOBile-YOLOv4 model as the testing benchmark. The comparison of losses under multiple models is shown in **Figure 9**.

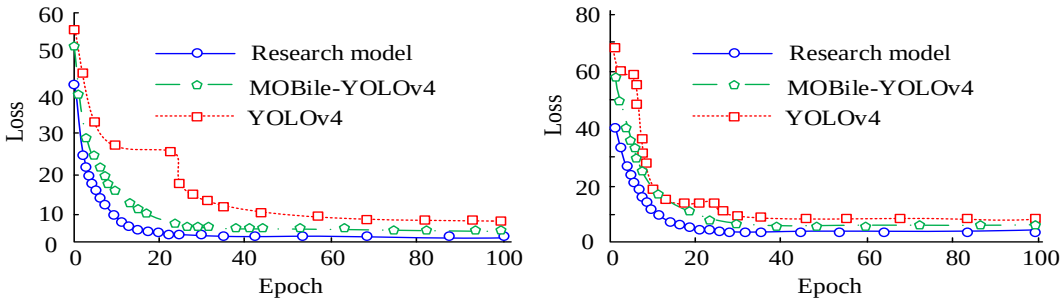

Figure 9. Comparison results of multiple model losses.

Figure 9a,b show the loss results for large-scale scenes and multi feature large-scale scenes, respectively. Compared to YOLOv4 and MOBile-YOLOv4, the research model had significant advantages in loss performance in both scenarios. In large-scale microbial scenarios, the loss value at model convergence was 0.12, while YOLOv4 and MOBile-YOLOv4 were 8.65 and 1.56, respectively. In more complex multi feature large-scale scenes, the loss value after convergence was 3.56, while MOBile-YOLOv4 and YOLOv4 were 8.65 and 11.25, respectively. The comparison outcomes of multiple model detection accuracy are denoted in **Figure 10**.

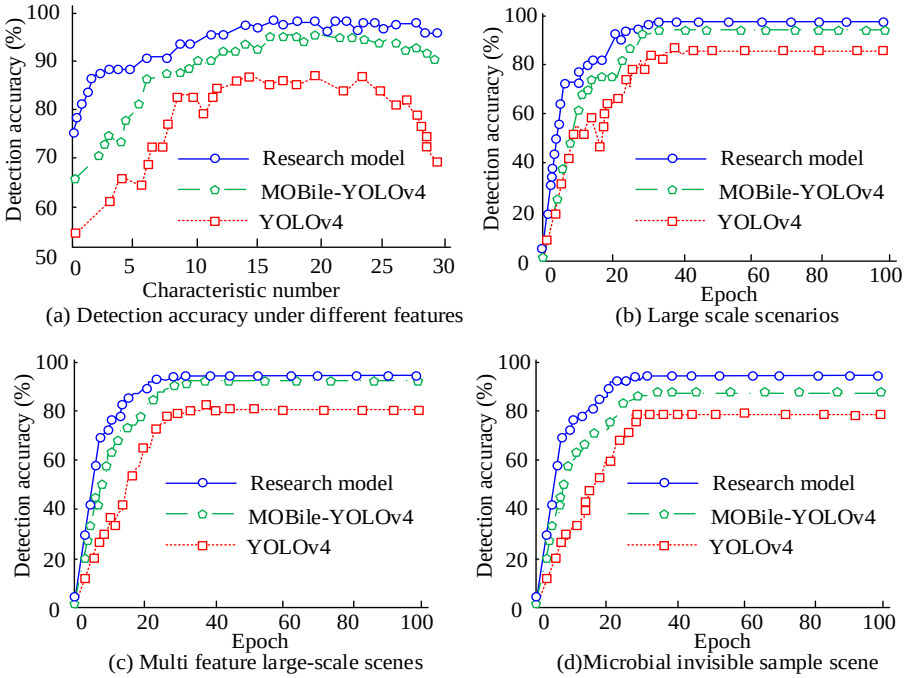

Figure 10. Comparison results of multiple model detection accuracy.

Three scenarios were selected for comparison, including multi feature scenes, large-scale scenes, and multi feature large-scale scenes as well as microbial invisible scenes, as shown in **Figure 10a–d**, respectively. In multi feature detection, when the number of microbial features is between 10 and 25, it is most advantageous for model detection. However, as the number of features exceeds 20, the accuracy of YOLOv4 detection significantly decreases. Overall, in complex feature environments, the research model shows a 12.25% and 19.52% improvement in detection performance compared to YOLOv4 and MOBile-YOLOv4. In the comparison of large-scale and multi feature large-scale scenes, the research model still has the highest detection accuracy, with detection accuracies of 98.65% and 97.58%, respectively, which are better than other models. In the scenario where microorganisms are not visible, the recognition performance of different technologies is studied. According to the test results, the overall performance of the research model is the best, with the highest recognition accuracy of 97.25% during convergence, while MOBile-YOLOv4 and YOLOv4 are 89.25% and 79.24%, respectively. The research model incorporates transfer learning strategies to enhance the capture and analysis of image edge contour details, enabling clearer detection of microorganisms in scenes that are invisible to ordinary people, resulting in better overall performance. Finally, common microorganisms in food were selected for detection, and microorganisms detection methods were added for comparison, as denoted in **Table 2**.

Table 2. Comparison of different microorganisms detection methods.

Type	Bioluminescence detection method (%)	YOLOv4 (%)	MOBile-YOLOv4 (%)	Research model (%)
Staphylococcus aureus	90.45	94.44	97.58	99.54
Salmonella	89.34	93.45	96.68	99.58
Escherichia coli	91.54	94.58	96.45	99.75
Shigella	90.24	93.45	95.54	98.54
Mould	91.54	94.84	94.56	99.24
Enterobacter	91.45	94.45	96.45	98.54
Erwinia	89.41	92.64	93.54	97.45
Escherichia coli	91.12	91.54	95.64	99.85
Flavobacterium	88.45	92.15	96.45	98.65
Enterococcus	89.56	93.45	97.45	99.85
Aeromonas	90.21	91.25	94.54	97.75

Table 2 shows the comparison results of different microorganisms detection methods. Among the four detection methods, the traditional luminescence detection method had the longest detection time and the lowest accuracy. In the comparison of Escherichia coli detection, the accuracy of bioluminescence detection method was 91.54%, YOLOv4 was 94.58%, and MOBile-YOLOv4 and the research model were 96.45% and 99.75%, respectively. It can be seen that computer vision-based microorganisms detection technology had better performance in food detection and met the requirements of food safety detection. In addition, the study compared the memory resource usage requirements of different technologies, as shown in **Table 3**.

Table 3. Comparison of memory resource utilization among different technologies.

Scene	YOLOv4 (%)	MOBile-YOLOv4 (%)	Research model (%)
Detection accuracy under different features	82.58	83.4	82.45
Large scale scenarios	83.15	84.45	83.45
Multi feature large-scale scenes	82.45	83.45	82.45
Microbial invisible sample scene	91.45	92.45	90.45
Multi microbial scene	87.45	89.45	86.45

According to the results in **Table 3**, the study selected five scenarios to analyze the memory resource occupation of different technologies. According to the results of the test volume, mobile-yolov4 has the highest memory resource occupation. For example, in the scene where microorganisms are invisible, the highest resource occupation rate is 92.45%, which is higher than 91.45% of yolov4 and 90.45% of the research model. In other scenarios, the research model has the same resource occupation as yoyov4, and is better than mobile-yoyov4. The main reason is that the research model optimizes the parameters of resnet 50 model, and strengthens the calculation of image edge details, so as to reduce the consumption of resources.

5. Conclusion

Traditional microorganisms detection technologies have low efficiency and average detection results. Therefore, a rapid microorganisms detection method combining computer vision and biotechnology was proposed in this study. Firstly, biotechnology was used to cultivate bacterial strains and create image sample data. Secondly, a microorganisms classification model was constructed based on ResNet50, and TL strategy was used for parameter learning optimization. Considering the difficulty of detecting large-scale complex scenes, a object detection model based on YOLOv4 was proposed, which replaced the feature extraction network with MobileNet and introduced an attention mechanism to enhance the attention of important features, thereby improving the detection effect on microorganisms. In the analysis of classification models, the loss of the research model, CBDT ResNet50, and CNN in low-density microorganisms detection were 0.48, 0.24, and 0.08, respectively, indicating better effectiveness of the research model. In the analysis of microorganisms detection effectiveness, the research model had a detection accuracy of 98.65% and 97.25% in the detection of high-density and low-density microorganisms, respectively, which was superior to other models. In the analysis of object detection models, common microorganisms detection effects were compared, and the accuracy of the research model in detecting Escherichia coli was 99.75%, while bioluminescence detection, YOLOv4, and MOBile-YOLOv4 were 91.54%, 94.58%, and 96.45%, respectively. The microorganisms detection technology proposed by the research had better adaptability and detection effectiveness in food safety detection. In summary, the technology proposed in this study has good application effects in practical scenarios. Compared with traditional manual detection, it has higher efficiency, multi scenario advantages, and higher detection accuracy, meeting the requirements of food safety and medical and health fields. However, this study also has limitations as it only analyzed common food

microorganisms. In the future, more microbial samples need to be added to improve the effectiveness of technical testing. At the same time, continue to optimize computer vision technology and enhance its application in more microbial recognition fields.

Funding: The research is supported by 2020 Fujian Province Young and Middle aged Teacher Education Research Project: Development and research of fermented beverage of red dragon fruit in northern Fujian (JAT201319).

Ethical approval: Not applicable.

Informed consent: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Ong K J, Johnston J, Datar I, Sewalt V, Holmes D, Shatkin J A. Food safety considerations and research priorities for the cultured meat and seafood industry. Comprehensive Reviews in Food Science and Food Safety, 2021, 20(6): 5421-5448.
2. Rani P, Kotwal S, Manhas J, Sharma V, Sharma S. Machine learning and deep learning based computational approaches in automatic microorganisms image recognition: methodologies, challenges, and developments. Archives of Computational Methods in Engineering, 2022, 29(3): 1801-1837.
3. Zhang J, Li C, Yin Y, Zhang J, Grzegorzek M. Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. Artificial Intelligence Review, 2023, 56(2): 1013-1070.
4. Xing G, Zhang W, Li N, Pu Q, Lin J M. Recent progress on microfluidic biosensors for rapid detection of pathogenic bacteria. Chinese Chemical Letters, 2022, 33(4): 1743-1751.
5. Xiao P, Liu J, Wang Z, Tao F, Yang L, Yuan G, Sun W, Zhang X. A color turn-on fluorescent probe for real-time detection of hydrogen sulfide and identification of food spoilage. Chemical Communications, 2021, 57(41): 5012-5015.
6. Guo Y, Girmatsion M, Li H W, Xie Y, Yao W, Qian H, Abraha B, Mahmud A. Rapid and ultrasensitive detection of food contaminants using surface-enhanced Raman spectroscopy-based methods. Critical Reviews in Food Science and Nutrition, 2021, 61(21): 3555-3568.
7. Yang M, Liu X, Luo Y, Pearlstein A J, Wang S, Dillow H, Reed K, Jia Z, Sharma A, Zhou B, Pearlstein D, Yu H, Zhang B. Machine learning-enabled non-destructive paper chromogenic array detection of multiplexed viable pathogens on food. Nature Food, 2021, 2(2): 110-117.
8. Zhang Y, Jiang H, Ye T, Juhas M. Deep learning for imaging and detection of microorganisms. Trends in Microbiology, 2021, 29(7): 569-572.
9. Firouz M S, Sardari H. Defect detection in fruit and vegetables by using machine vision systems and image processing. Food Engineering Reviews, 2022, 14(3): 353-379.
10. Martín J P, Alvarado-Capó Y, Morales R O, Pichardo T, López A A. Count of bacteria and yeast in microbial bioproduct using digital image processing. ITEGAM-JETIA, 2021, 7(32): 12-22.
11. Harakannanavar S S, Rudagi J M, Puranikmath V I, Siddiqua A, Pramodhini R. Plant leaf disease detection using computer vision and machine learning algorithms. Global Transitions Proceedings, 2022, 3(1): 305-310.
12. Liang N, Sun S, Zhang C, He Y, Qiu Z. Advances in infrared spectroscopy combined with artificial neural network for the authentication and traceability of food. Critical Reviews in Food Science and Nutrition, 2022, 62(11): 2963-2984.
13. Jayan H, Pu H, Sun D W. Recent developments in Raman spectral analysis of microbial single cells: Techniques and applications. Critical Reviews in Food Science and Nutrition, 2022, 62(16): 4294-4308.
14. Mishra G, Panda B K, Ramirez W A, Jung H, Singh C B, Lee S H, Lee I. Research advancements in optical imaging and spectroscopic techniques for nondestructive detection of mold infection and mycotoxins in cereal grains and nuts. Comprehensive Reviews in Food Science and Food Safety, 2021, 20(5): 4612-4651.

15. Wang J, Jiang Z, Wei Y, Wang W, Wang F, Yang Y, Song H, Yuan Q. Multiplexed identification of bacterial biofilm infections based on machine-learning-aided lanthanide encoding. ACS Nano, 2022, 16(2): 3300-3310.
16. Khaled A Y, Parrish C A, Adedeji A. Emerging nondestructive approaches for meat quality and safety evaluation-A review. Comprehensive Reviews in Food Science and Food Safety, 2021, 20(4): 3438-3463.
17. Sharma S, Gahlawat V K, Rahul K, Mor RS, Malik M. Sustainable innovations in the food industry through artificial intelligence and big data analytics. Logistics-Basel, 2021, 5(4): 66.
18. Mokayed H, Quan T Z, Alkhaled L, Sivakumar V. Real-time human detection and counting system using deep learning computer vision techniques. Artificial Intelligence and Applications, 2023, 1(4): 221-229.
19. Simon K, Vicent M, Addah K, Bamutura D, Atwiine B, Nanjebe D, Mukama A O. Comparison of deep learning techniques in detection of sickle cell disease. Artificial Intelligence and Applications, 2023, 1(4): 252-259.
20. Justé A, Thomma B P, Lievens B. Recent advances in molecular techniques to study microbial communities in food-associated matrices and processes. Food Microbiology, 2008, 25(6): 745-761.
21. Hernández Medina R., Kutuzova S., Nielsen K N, Johansen J, Hansen L H, Nielsen M, Rasmussen S. Machine learning and deep learning applications in microbiome research. ISME Communications, 2022, 2(1): 98.
22. Moreno-Indias I, Lahti L, Nedyalkova M, et al. Statistical and machine learning techniques in human microbiome studies: Contemporary challenges and solutions. Frontiers in Microbiology, 2021, 12: 635781.
23. Marcos-Zambrano L J, Karaduzovic-Hadziabdic K, Loncar Turukalo T, et al. Applications of machine learning in human microbiome studies: A review on feature selection, biomarker identification, disease prediction and treatment. Frontiers in Microbiology, 2021, 12: 634511.
24. Marcos-Zambrano L J, López-Molina V M, Bakir-Gungor B, et al. A toolbox of machine learning software to support microbiome analysis. Frontiers in Microbiology, 2023, 14: 1250806.