A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches

Pingli Ma[1] · Chen Li[1] · Md Mamunur Rahaman[1] · Yudong Yao[2] · Jiawei Zhang[1] · Shuojia Zou[1] · Xin Zhao[3] · Marcin Grzegorzek[4]

Abstract

Microorganisms play a vital role in human life. Therefore, microorganism detection is of great significance to human beings. However, the traditional manual microscopic detection methods have the disadvantages of long detection cycle, low detection accuracy in large orders, and great difficulty in detecting uncommon microorganisms. Therefore, it is meaningful to apply computer image analysis technology to the field of microorganism detection. Computer image analysis can realize high-precision and high-efficiency detection of microorganisms. In this review, first,we analyse the existing microorganism detection methods in chronological order, from traditional image processing and traditional machine learning to deep learning methods. Then, we analyze and summarize these existing methods and introduce some potential methods, including visual transformers. In the end, the future development direction and challenges of microorganism detection are discussed. In general, we have summarized 142 related technical papers from 1985 to the present. This review will help researchers have a more comprehensive understanding of the development process, research status, and future trends in the field of microorganism detection and provide a reference for researchers in other fields.

✉ Chen Li
lichen201096@hotmail.com

1   Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

2   Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA

3   School of Resources and Civil Engineering, Northeastern University, Shenyang, China

4   Biomedical Information College, University of Luebeck, Luebeck, Germany

1 Introduction

1.1 Knowledge of microorganism

Microorganisms have a unique ability to adapt to extreme conditions. They are found in every environment imaginable. Microorganisms refer to tiny organisms with independent living functions, which means they can absorb energy, grow and reproduce by themselves. The modern classification of microorganisms are in a three-domain system. The system consists of Archaea, Eucarya, and Bacteria (Pepper et al. 2011). Bacteria, along with actinomycetes and cyanobacteria (blue-green algae) belongs to the prokaryotes. Eukaryotes or eukarya includes fungi protozoa, algae, plant, and tiny animal. Viruses are obligate intracellular parasites that belong to neither of these two groups (Bitton 2005). Microorganisms play a very important role in human life, for better or worse. For example, the presence of plant growth-promoting rhizobacteria can encourage beneficial effects on plant health and growth. It can also suppress disease-causing microbes (Babalola 2010). Deleterious rhizobacteria can inhibit plants growth through the production of phytotoxins (Nehl et al. 1997). Lactic acid bacteria can affect humans in many aspects, such as the regulation of gastrointestinal flora, the normal operation of human metabolism, and the inhibition of the reproduction of harmful bacteria (Masood et al. 2011). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can cause fever, discomfort, dry cough, shortness of breath and even death in severe case (Hui et al. 2019). In summary, the work of microorganism analysis is of great significance for the human being.

Observation under a microscope is a common and important method in microorganism analysis. A stereo scan electron microscopy uses microorganism in soil analysis (Gray 1967). In Daley and Hobbie (1975), a modified epifluorscence technique based on a microscope is used for aquatic bacteria counting. In Collins et al. (1993), environmental scanning electron microscopy is used for microorganism analysis. However, those microscopic methods have some disadvantages. Firstly, there are many types of microorganisms. Estimated in Locey and Lennon (2016), Earth is inhabited by 10^{11}–10^{12} microbial species. Thus the knowledge of experts is always inadequate. When experts use this method for microorganism analysis, they often have to consult a great deal of literature. Secondly, the training cycle for researchers and overall detection time need much time. For example, the counting of phytoplankton with traditional microscopic method requires much time. Besides, an operator needs to master a great deal of professional knowledge (Embleton et al. 2003). Thirdly, the microorganisms to be analyzed are often of large orders of magnitude. It is difficult for microscopic methods to deal with the analysis problem with a large amount of data. Large sample size will affect the analysis accuracy of operators (Van et al. 2002). Because of these disadvantages of the microscopic method, we need a more efficient method for microorganism analysis. For example, computer image analysis is a feasible method.

1.2 Motivation

Computer image analysis is part of both computer vision and image processing. Typically, image analysis is used for gaining insight into the raw image and extracting the information we need (Umbaugh 2005). In terms of microorganism analysis, using image analysis has many advantages over the microscopic method. Firstly, image analysis does not care

about the number of microbial species. Computers can store and remember more information about different kinds of microorganisms than experts. The Ribosomal Database Project stores 2460 different species. As information is entered, the number will continue to increase (Maidak et al. 2000). Secondly, image analysis takes less time. With the help of image analysis, phytoplankton analysis just needs 30–40 min (Embleton et al. 2003). Thirdly, for image analysis, the task of microorganism analysis of large orders of magnitude does not pose any difficulties. Fourthly, the operation of image analysis is simple. When using image analysis, necessary operations usually only include import data and a few simple steps, such as the Center for Microbial Ecology Image Analysis System (Dazzo and Niccum 2015). Based on the above reasons, it is feasible to apply computer-aided image analysis in microorganism analysis.

At present, many people are conducting researches in this field, such as microorganism segmentation (Kulwa et al. 2019), microorganism clustering (Li et al. 2020), microorganism classification (Li et al. 2019), microorganism counting (Li et al. 2021), etc. However, there is no specific review of microorganism detection. To have a comprehensive overview of the existing microorganism detection research, we discuss many materials. According to our search, some reviews involve microorganism detection. In Benfield et al. (2007), the development of plankton-imaging systems and advances in extracting information from image data sets timely are summarized, where twelve papers involve microorganism detection among all 56 references. In Schaap et al. (2012), a review of the current status of lab on a chip technologies in the context of algae detection and monitoring is presented. In this work, there are twenty four papers related to algae detection mentioned among all 81 references. In Gopinath et al. (2014), a review for existing bacterial detection methods is proposed, where methods from manual microscope detection to smartphone-based detection are analysed. Twelve papers in this survey are related to bacterial detection among all 126 references. One review focuses on using computer image analysis to examine the microscopic objects (Puchkov 2016). In this work, twenty-two papers are related to microorganism detection among all 102 references. In Li et al. (2019), the development history of microorganism classification using content-based microscopic image analysis approaches is reviewed, where twenty-four papers relate to microorganism detection among all 317 references. In Zhou et al. (2020), the development of diatom testing over the decades is reviewed and a new method of deep learning for diatom detection is discussed. In this work, two papers are related to diatom detection among all 34 references. Hence, none of these surveys conducts a comprehensive study on object detection in microorganism image analysis. In fact, the analysis of microorganism detection methods can facilitate development in other microorganism image analysis fields, such as microorganism image classification (Zhao et al. 2021; Kulwa et al. 2021; Li et al. 2021; Xu et al. 2020; Li et al. 2019; Kosov et al. 2018; Li et al. 2016, 2015), microorganism image segmentation (Zhang et al. 2021a, b, 2020; Kulwa et al. 2019) and microorganism image retrieval (Zou et al. 2016, 2017).

As shown in Fig. 1, for microorganism detection, the early work is mainly based on classical image processing methods, such as image segmentation (Fukuda and Hasegawa 1989) and binarization (Bloem et al. 1995). With the development of the field of machine learning, researches used traditional machine learning methods appear gradually, such as artificial neural network (ANN) (Widmer et al. 2005), support vector machine (SVM) (Lenseigne et al. 2007) and genetic algorithm-neural network (GA-NN) Osman et al. (2010). In recent years, deep learning-based methods become very popular, such as faster region-convolutional neural networks (Faster R-CNN) (Viet et al. 2019) and Mask R-CNN (Ruiz-Santaquiteria et al. 2020). In addition, many new methods can be applied in
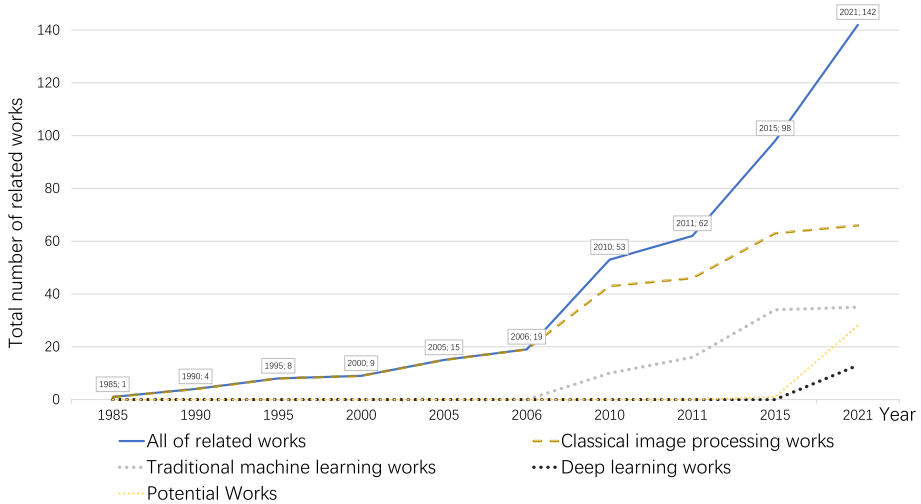
Fig. 1 The total number of related works on microorganism detection

microorganism detection, such as single shot detector (SSD) (Liu et al. 2016). Moreover, visual transformer-based methods have a more robust global information representation capability than the current famous CNNs, which means they can eliminate the problem of microorganism structure described in the complete image. Therefore, the field of visual transformer has developed rapidly in recent years. In fact, visual transformer has been applied to cell detection with satisfactory results (Aubreville et al. 2017). However, visual transformer have not yet been applied to microorganism detection, although visual transformer based methods show great potential for application in microorganism detection, such as squeezeand-excitation (SE) (Hu et al. 2018), Vision Transformer-Faster RCNN (ViT-FRCNN) (Beal et al. 2020). The number of related works is rising steadily. The data in Fig. 1 shows that microorganism object detection has a good development trend and a great development potential. Thus we decide writing a survey of object detection technologies for microorganism image analysis, where we summarize about 142 related works for a comprehensive survey. Methods mentioned in these works are mainly grouped into classical image processing based methods, traditional machine learning based methods, deep learning based methods and potential methods, as shown in Fig. 2.

1.3 Paper retrieval and screening

We specifically searched for papers in the field of microbiology. Therefore, according to the definition of microorganisms (tiny organisms with independent living), some excellent articles are not analyzed and discussed in this article due to the non-compliant research fields, such as (Aubreville et al. 2017) for histology cell, (Haoyuan et al. 2021) for gastric histopathology image, (Liu et al. 2021) for cytopathology cell and (Hu et al. 2022) for gastric histopathology image.

The retrieval and screening flow of papers involved in this review is shown in Fig. 3. In retrieving stage, we first search for random combinations of the following phrases in Google Scholar: microbe, detection, recognition, classification, image analysis, image processing, where 575 papers are collected. Then, based on the references of collected
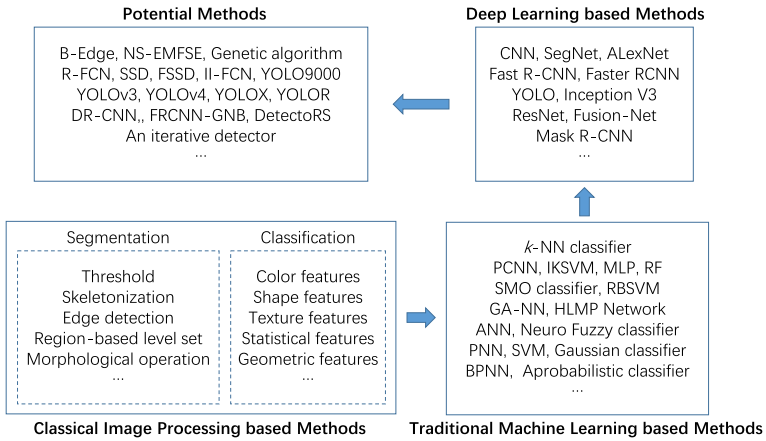
Potential Methods

B-Edge, NS-EMFSE, Genetic algorithm
R-FCN, SSD, FSSD, II-FCN, YOLO9000
YOLOv3, YOLOv4, YOLOX, YOLOR
DR-CNN,, FRCNN-GNB, DetectoRS
An iterative detector
…

Deep Learning based Methods

CNN, SegNet, ALexNet
Fast R-CNN, Faster RCNN
YOLO, Inception V3
ResNet, Fusion-Net
Mask R-CNN
…

Segmentation

Threshold
Skeletonization
Edge detection
Region-based level set
Morphological operation
…

Classification

Color features
Shape features
Texture features
Statistical features
Geometric features
…

k-NN classifier
PCNN, IKSVM, MLP, RF
SMO classifier, RBSVM
GA-NN, HLMP Network
ANN, Neuro Fuzzy classifier
PNN, SVM, Gaussian classifier
BPNN, Aprobabilistic classifier
…

Classical Image Processing based Methods

Traditional Machine Learning based Methods

Fig. 2 Microorganism image detection methods mentioned in this survey

Paper retrieval

First round search

Based on key words: 575 papers

Second round search

Based on references of retrieved papers: 1326 papers

Initial dataset: 1901 papers

Paper screening

First round screening

Based on duplicates: removing 96 papers

Second round screening

Based on title and abstract: removing 1328 papers

Third round screening

Based on full text: removing 335 papers

114 papers
(Otsu, SVM, R-CNN …)

Existing methods

Remaining 142 papers
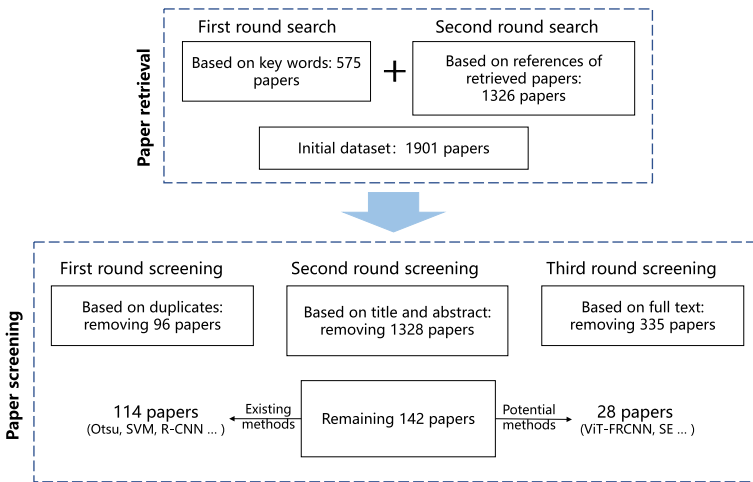
Potential methods

28 papers
(ViT-FRCNN, SE …)

Fig. 3 The retrieval and screening flow of involved papers

papers, we accumulated extra 1326 papers. Therefore, a total 1901 papers are obtained in the retrieving stage. In screening stage, we mainly removing papers by three steps. Firstly, we screen paper by checking whether the paper is duplicated. 96 papers are removed. Among retained papers, 1325 papers' abstract or title are not related to our review and then removed. At last, 335 of remaining papers are removed after reading through them carefully. In brief, 142 papers are remained and analysed in our review. Among them, 65 papers are related to classical image processing; 34 papers are related to traditional machine learning; 13 papers are related to deep learning; one is related to both classical image processing and traditional machine learning; one is related to both traditional machine learning and deep learning; 28 papers are related to potential methods, including visual transformer-based methods.

1.4 Structure of this review

In this paper, microorganism detection is summarized according to different detection methods. Firstly, the existing microorganism detection methods are classified. Then, each class is described and analyzed respectively. In addition, the research motivation, basic knowledge of object detection and evaluation criteria of different methods are introduced. For some critical methods, flow charts or sample diagrams are shown to deepen understanding.

The rest of this review is structured as follows: In Sect. 2, we have introduced the basic knowledge of object detection and some commonly used evaluation indicators. Then, in Sects. 3, 4 and 5 we have introduced classical image processing-based methods, traditional machine learning-based methods and deep learning-based methods,respectively. After, the three categories of methods mentioned above are summarized in Sect. 6. In addition, potential methods are analysed. Finally, Sect. 7 summarizes the whole work with potential future work. With the structure, relevant researchers can have a better understand of the related works.

2 Overview of microorganism detection

Prior to an overview of microorganism detection approaches, we review the basic knowledge of object detection. Next, depending on the form of the test results detection results, the history of microorganism detection is grouped into the early phase and current phase. We will outline the early and current phases of microorganism detection development in chronological order.

2.1 Basic knowledge of object detection

Object detection is a long-term and challenging task in computer vision (Fischler and Elschlager 1973). As one of the primary tasks of computer vision, the ultimate goal of object detection is to give the classes and locations of objects. Therefore, object detection is the basis for many other computer vision tasks, such as object tracking (Son et al. 2017) and scene understanding (Pan et al. 2017). Detection methods are evolving all the time since object detection is proposed. In this section, we comprehensively summarize the development of object detection methods over the past twenty years and group object detection methods into traditional manual features-based methods and deep learning-based methods (Zou et al. 2019). Among them, deep learning-based methods include one-stage methods and two-stage methods according to different processing flows. Figure 4 shows the development of detection methods over the past two decades. Methods involved in Fig. 4 are as follows: Viola-Jones(VJ) (Viola and Jones 2001), Histograms of Oriented Griented gradients(HOG) (Dalal et al. 2005), Deformable Parts Model (DPM) (Felzenszwalb et al. 2008), R-CNN (Girshick et al. 2014), Spatial Pyramid Pooling Networks(SPPNet) (He et al. 2015), Fast R-CNN (Girshick 2015), Faster R-CNN (Ren et al. 2016), Pyramid Networks (Lin et al. 2017), Mask R-CNN (He et al. 2017), OverFeat (Sermanet et al. 2013), You Only Look Once (YOLO) (Redmon et al. 2016), Single Shot MultiBox Detector (SSD) (Liu et al. 2016), YOLO9000 (Redmon et al. 2017), Retina-Net (Lin et al. 2017), YOLOv3 (Redmon and Farhadi 2018),
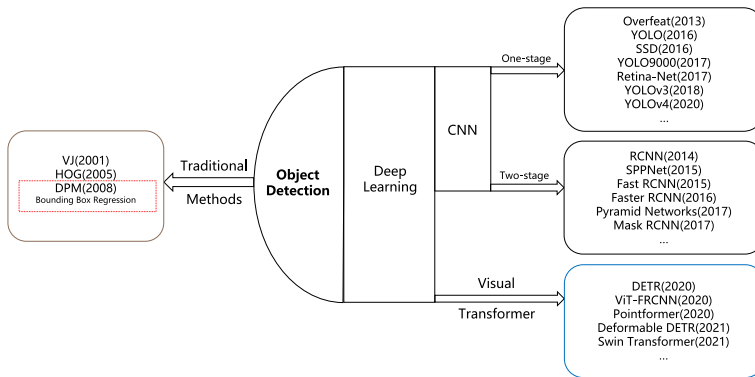
Fig. 4 The development of detection methods over the past two decades

YOLOv4 (Bochkovskiy et al. 2020), Vision Transformer-Faster RCNN (ViT-FRCNN) (Beal et al. 2020), DEtection TRansformer (DETR) (Carion et al. 2020), Pointformer (Pan et al. 2020), Deformable DETR (Zhu et al. 2021) and Swin Transformer (Liu et al. 2021).

2.2 The early phase of microorganism detection

In the early phase of microorganism detection, the primary purpose is to determine whether the object exists or not. At this phase, microorganism detection is closely related to microorganism segmentation. Figure 5 shows an example of segmentation-based detection. In Bloem et al. (1995), the detection of soil bacterium is based on the combination of segmented binary and grayscale. In (Rachna and Swamy 2013), *k*-means clustering and Otsu threshold are respectively employed to segment mycobacterium tuberculosis images for mycobacterium tuberculosis detection. In addition, microorganism detection is also closely related to classification. In Lenseigne et al. (2007), the automatic detection of mycobacterium tuberculosis is based on SVM classifier. In Veri-kas et al. (2014), SVM as well as random forest (RF) classifiers is used for distinguishing prorocentrum minimum (P. minimum) cells and other objects.
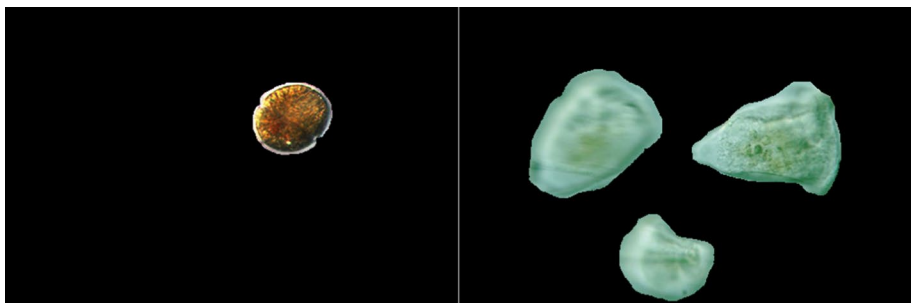

Fig. 5 Examples of segmentation-based detection

2.3 The current phase of microorganism detection

Unlike the early phase of microorganism detection, the current phase of microorganism detection precisely gives the classes and locations of objects. At this stage, the detection results include several bounding boxes surrounding each detected objects with respective class labels. Microorganism detection at this stage is mainly based on deep learning methods (LeCun et al. 2015). An example of a deep learning-based detection result is shown in Fig. 6.

Deep learning has a long history of development. In 1943, the emergence of McCulloch-Pitts (MCP) model, which is also regarded as original version of the artificial neural model (ANN), means the initial realization of deep learning (McCulloch and Pitts 1943). After MCP, deep learning field has a long period of stable development until the problem caused by hardware limitations is resolved. After 2014, thanks to the powerful computing power of graphics processing unit (GPU) and implementation of CNN models, deep learning has been able to develop rapidly, especially in object detection task. In fact, Faster R-CNN, a typical model of two-stage object detector not proposed until 2016. As for YOLO, a typical model of one-stage object detector, it is proposed even five months later than Faster R-CNN. It is only after the relative maturity of deep learning techniques that it is applied to microorganism detection. Therefore, while deep learning is widely used for object detection, few research studies are based on deep learning for microorganism detection. However, it can be seen from Fig. 1 that deep learning in the field of microorganism detection has developed rapidly, although it started late. After consulting a lot of materials, we have found 13 researches that are related to deep learning. In Hung and Carpenter (2017), Faster R-CNN is firstly applied for cell detection in brightfield microscopy images of malaria-infected blood. In Ruiz-Santaquiteria et al. (2020), Mask R-CNN is for the first time used for detecting diatoms in images, where include many diatom shells.

2.4 Evaluation criteria

In order to evaluate the effectiveness of microorganism detectors, some proper criteria are given in different related work. It is observed that there is no widely accepted evaluation criteria for evaluating detection performance. Different criteria are applied in different research scenarios. For detection work in counting, the main criterion is the correlation with the expert counting results or real number, such as Ogawa et al.
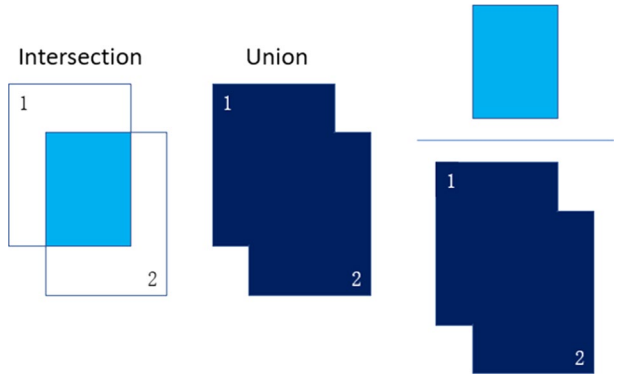

Fig. 6 Example of deep learning-based detection

Table 1 Explanatory table of TP, FN, FP, and TN

True label	Predicted label	
	Yes	No
Yes	TP	FN
No	FP	TN

Fig. 7 Illustration of the calculation of IoU

(2005), Hung and Carpenter (2017). When using this kind of criterion, the expert counting results or actual number is regarded as the baseline. Therefore, the closer the detection results are to the reference baseline, the better. In Jan et al. (2015), the effectiveness of detection result for tubercle bacillus(TB) is evaluated by the number of objects in sample image and the number of detected objects, shown in Eq. 1.

$$Accuracy = \frac{\text{No.objects in sample image}}{\text{No.detected TB bacteria}} \tag{1}$$

For detection work in classification, accuracy, sensitivity, specificity and precision are widely used as the criteria. To better explain how to calculate these evaluation criteria, we illustrate four fundamental concepts in Table 1. Positive and negative represent the judgement results of the model. True and false indicate whether the model's decision is correct or not. By combining positive, negative, true and false, we can get true positive (TP), false positive (FP), true negative (TN) and false negative (FN). In addition, to judge whether the actual detection result belongs to TP, FP, TN or FN, not only the detection category but also the Intersection over Union (IoU) must be considered. In practice, IoU is used to evaluate the distance between the output box and the ground-truth. The calculation method of IoU is shown in Fig. 7, where box 1 and box 2 respectively represent real box and prediction result, the light blue area and the dark blue area represent the intersection and union of the predicted box and the actual box in turn. Finally, IoU is calculated by intersection ratio union.

Based on the four criteria in Table 1, the calculation methods of accuracy, sensitivity, specificity and precision are shown in Table 2. Some studies only used accuracy (Hiremath and Bannigidad 2010) or sensitivity (Verikas et al. 2012) as evaluation metrics. However, few works considered multiple of them, such as sensitivity and specificity (Shah et al. 2016), and accuracy, sensitivity and specificity (Osman et al. 2010).

Table 2 Formula of accuracy, sensitivity, precision and specificity

	Accuracy	Sensitivity	Specificity	Precision
Formula	$\frac{TP+TN}{TP+TN+FP+FP}$	$\frac{TP}{TP+FN}$	$\frac{TN}{FP+TN}$	$\frac{TP}{TP+FP}$

In addition to the common criteria mentioned above, F-score, mean average precision(mAP), true accept rate (TAR), false accept rate (FAR), false reject rate (FRR) and detection accuracy are also used. Table 3 shows the mathematical expression of them. For the calculation of F-score, recall is the same as sensitive mentioned before. For mAP, AP means the mean of precision-values for each class and m means how many classes detected. The value of AP is determined by the precision-sensitivity curve.

2.5 Summary

Microorganism detection methods have changed dramatically over the decades. In terms of expressions, microorganism detection develops from simple counting to the realization of localization and classification. In implementation techniques, microorganism detection develops from direct observation with a microscope to more efficient methods using modern image processing techniques. With the continuous improvement of detection methods, the corresponding evaluation criteria are also designed.

3 Classical image processing based methods

In this section, classical image processing based methods for microorganism detection are grouped into segmentation based methods and classification based methods respectively introduced in Sects. 3.1, 3.2. After that, a concise table is prepared in Sect. 3.3. In Table 4, the publication date, literature links, research objects and its domains, main methods and evaluation indicators of related works are presented.

3.1 Segmentation based methods

In this subsection, related works that used segmentation methods for microorganism detection are introduced.

In Sieracki et al. (1985), for alleviating subjectivity of manual counting, a novel method based on a epifluorescence microscope and image analyzer is proposed. The first step is image enhancement, including setting small bright dots next to each detected object and bright enhancement covering the entire detected area of each object. The second step is using a light pen manually for distinguishing the object from other regions. The last step is object counting with the general analysis software provided by Artek Systems Corp. The effectiveness

Table 3 Formula of mAP, TAR, FAR , FPR and F-score

	mAP	TAR	FAR	FRR	F-score
Formula	$\frac{1}{n} * \sum_{i=1}^{n} AP_i$	$\frac{TA}{TA+FA+FR}$	$\frac{FA}{TA+FA+FR}$	$\frac{FR}{TA+FA+FR}$	$\left(1+\beta^2\right) * \frac{Precision*Recall}{\beta^2*Precision+Recall}$

of proposed system is analysed by comparing counts of image analysis with visual counts. Experiment result shows that the proposed system can achieve the same accuracy with visual counts in bacterial detection.

In order to analyze the characterization of mycelial morphology, in Adams and Thomas (1988), a semi-automated image analysis is introduced. The first step is preprocessing by a mean operator and an edge enhancement operator for denoising and sharpening the image. The second step is segmentation and skeletonization. After that, the light pen is used to identify the branches manually. Finally the length of the main hyphae is determine based on the software supplied with the Magiscan 2A. The experiment is performed on five samples. The result shows that compared with the traditional digitizing-table method, semi-automatic image analysis has the advantages of short time-consuming and high convenience.

In Packer and Thomas (1990), a fully automatic system is developed for morphological measurements on filamentous microorganisms, in which speed is gained. The first step is to binarize the image based on the preset gray value. The second step is eliminating objects in processed image other than microorganisms such as dust, media particles by a circularity test. Each objects in the resulting image is then skeletonized. Subsequently, the processed binary image is divided into an image containing measurable microorganisms and an image containing clumped (or aggregated) material. Finally, the processed image are measured. Experiment is performed on eight samples of about 100 microorganisms. The result shows that the fully automatic method is marginally faster than the manual image analysis method and much faster than the digitizing method.

In Masuko et al. (1991), a novel anultra-high-sensitivity TV camera based method for enumeration of bacteria is proposed. The first step is using a anultra-high-sensitivity TV camera for enumerating small objects that emit light in a photon-counting image. After colony growth, a luminous image of the membrane filter is obtained. Finally, the accuracy of counting is judged by comparing the photon-counting image with the luminous image of the membrane filter. Experiment result shows that the for proposed algorithm is reliable for detecting single bacteria with light.

In Bloem et al. (1995), for counting cell numbers and calculating some geometric parameters of bacteria in soil smears, a fully automatic image processing method is proposed. Firstly, convolution, morphological erosion and dilation to remove noise are used for smoothing grey images. Holes in the image are then filled for background equalization. By applying two top hat transforms, the background is eliminated. At last, an image sharpening operation is done for better detention. Particles that have a higher threshold than a fixed one are detected. In addition, the cells number is determined by number of detected particles with maximum gray scale. The proposed method has high efficiency for detecting about 1500 cells within 30 min. In addition, the detection result of proposed method is similar to the manual detection result.

In Baillieul and Scheunders (1998), an image analysis system is described for measuring the average velocity of many simultaneously moving similar objects. First, the operator manually locates an active window on the image to focus on the area of interest. The second step is setting the object to one and the background to zero based on a fixed threshold value. At last, by collecting the trajectories in real-time over a series of frames, average velocity of the objects are calculated. Three groups of 25 daphnids are tested on this experiment. Result shows that the system is a useful tool for the objective quantification of velocity.

In Wang et al. (2003a), a statistical, non-parametric framework based method is proposed for three-dimensional (3-D) object detection and labelling. The first step is extracting the nuclei from the observed image. In this step, the image is segmented into regular

non overlapped cubes. Each cube is then inspected based on defined criteria. The cubes identified as the nucleus are setting to white. At last, all white cubes in the same cell are merged into a whole by an iterative merging algorithm. Figure 8 indicates the results obtained based on proposed method. Experiment suggests that proposed method achieves a fast cell detection on prepared data within 148 seconds.

In Wang et al. (2003b), a low computational complexity and robust method for 3-D biological object detection and labelling is described. The first step is extracting the nuclei from the observed image. In this step, the image is divided into regular non overlapped cubes. Each cube is then inspected based on defined criteria. The cube identified as the nucleus is setting to white. At last, all white cubes in the same cell are merged into a whole by an iterative merging algorithm. Several experiments are performed on real image data to demonstrate the applicability. 439 nuclei are detected in the experiments.

In Qing et al. (2006), a method for identifying and counting algae microscopic color images is proposed. This method sequentially performs the following operations on the image: image smoothing by median filtering, threshold segmentation based on the hue saturation and intensity (HSI) color space characteristics of the image, seed filling, noise elimination based on the noise region threshold segmentation, refinement and the final step of the total calculation quantity. The experiment compares the performance of the proposed
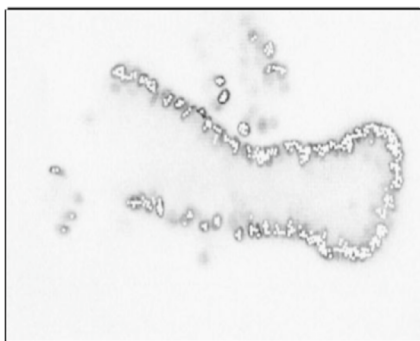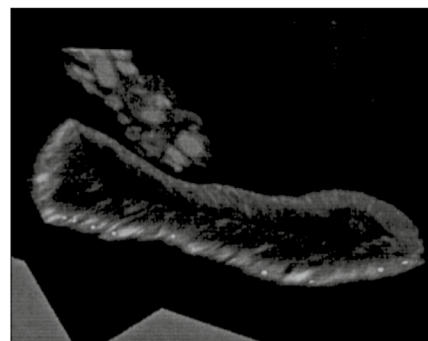

(a) Germs detection
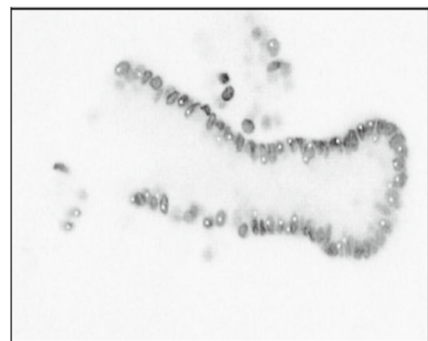
(b) Nuclei labelling
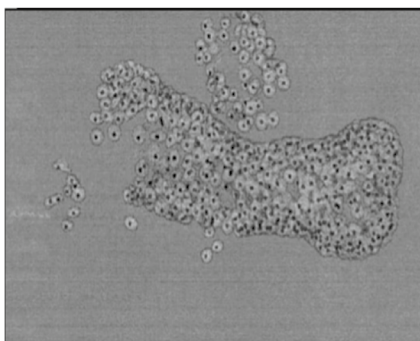
(c) Volume rending of the labelling results

(d) Dynamic inspection of volume rendered images with objects incorporated

Fig. 8 The overview of proposed method mentioned in Wang et al. (2003a). The figure corresponds to Fig. 8 in the original paper

system with manual statistics on 40 images of algae cells containing Chlamydomonas and Chlamydomonas bicuspidata. The result shows that compared with manual recognition method, the recognition rate of proposed system is greater than 90%, indicating that this method is feasible.

For improving the accuracy of traditional image segmentation and image enhancement algorithms, in Li and Chen (2007), an image edge detection method based on histogram equalization and soft mathematical morphology is proposed. This method is on the basis of combining traditional image enhancement and image segmentation. The first step is pre-processing, including histogram equalization and histogram specification. The second step is flexible morphological operations, including corrosion and expansion. The last step is edge detection based on the degree of gray value change. Based on this method, the detection accuracy can be improved and the edge details can be protected well.

In Costa et al. (2008), in order to reduce the labor workload of clinician while diagnosing tuberculosis and the cost of the patient, an automatic detection algorithm is proposed for TB. Firstly, a difference image (R channel minus G channel) from red green and blue (RGB) color format is obtained. Secondly, a global adaptive threshold operation is applied for TB segmentation. At last, several filtering operations are used for separating the bacilli from artifacts. About 50 images are analysed in the experiment, and the number recognized by experts is regarded as a reference standard. Experiment result shows that the proposed algorithm yields the highest sensitivity of 76.65% and the best FP rate of 12%.

In Zhang et al. (2008), for improving the counting efficiency of bacterial colonies, a fully automatic and cost-effective bacterial colony counter is introduced. First, original images are classified into color images and non-color images. For all images, Otsu threshold is applied for segmenting foreground object region, which is named as the dish/plate area. After that, for color image, Otsu threshold is applied for further objects segmenting from dish/plate area. Watershed algorithm is employed for final segmentation of object with more than one colony. For non-color image, the further objects segmenting from dish/plate area is based on a statistic method. The final detection number is calculated by adding the number detected in the color image to the number detected in the non-color image. For testing experiment, 100 chromatic and achromatic images are prepared. In addition, Otsu's method is employed for a controlled experiment. Result shows that the presented algorithm has better performance with a satisfaction rate of 96% compared to Otsu's method.

In Wang et al. (2008), an automatic method is introduced to track motile microorganism. Microorganisms in motion are tracked through the movement of the XY platform. The main image processing method used in this work is multiple thresholds based object recognition. For the binary image, the centroid of the target mask is regard as the coordinates to calculate the speed and direction of the target movement in the time series image. The tracking control system is tested on the unicellular photosynthetic alga Chlamydomonas reinhardtii. Experiment result shows that combined with proposed microorganism detection method, the system can continuously track cells in motion for up to 300 s.

In Fernandez-Canque et al. (2006) and Fernandez et al. (2008), an automatic machine vision based method is presented for micro-organism oocysts detection, which can not only provide satisfactory detection results, but also require very little time. Two detection methods are mentioned in this system. The first method is threshold twice. Firstly, the brighter and darker components of the image are respectively processed by threshold. After that, an or operation and denoising are performed on two images processed. At last, the objects are detected by employing Danielsson circle detection method. The other method is based on independent red color space and green color space. Based on the green plane, the wall and cell nucleus of Cryptosporidium are detected. Based

on the red plane, only the nucleus are detected. The number of nuclei is obtained from the red image. The number of cell walls is obtained by subtracting the two. Experiment result shows that the proposed method achieve a successful detection rate of 100%.

In Coltelli et al. (2008), for measuring the velocity of moving objects in microscopic images, a novel digital system is proposed. The first step is a subtraction operation between two successive frames. Based on this step, moving objects in video are detected while stationary objects are eliminated. The difference images are then stored. The third step is automatic labelization of the cells moving. At last, the variation of area is calculated and regarded as a measure of speed. 10-images time sequence (400 msec) are used for test. The test data shows that the speed of *Dunaliella* cells is about 100 μ/s, which is similar to the speed of *Dunaliella* cells of 100 μ/s reported in previous studies Rose (1974). This means the proposed system can accurately measure the speed of moving object.

For providing the doctors a faster way to diagnose tuberculosis and allowing patients to spend less, in Sotaquira et al. (2009), a image processing techniques based method is described. This method can perform bacterial segmentation and cluster detection for TB diagnosis purposes. First, color space of the original image in RGB is converted into YCbCr and Lab color spaces. Then Cr plane of YCbCr color space and a plane of Lab color space are then respectively segmented by threshold calculated. At last, two segmented images are combined for final detection. 1400 images of 14 patients are used for testing this method, which takes the ratio of correctly segmented objects to the total number of objects detected in the image as standard. Experiment results suggest that method proposed in this work yields a good detection performance with an average efficiency of 96.3%.

In Vallotton et al. (2010), a novel detection method is proposed for characterizing bacteria, that is from high resolution phase contrast image. The first step is boundary detection by Marr Hildreth edge detector. The second step is selecting features for detecting splitting, including the image contrast at a particular candidate point, a constriction at the septum, feature provided by a probabilistic model and angular feature. After that, the normalization operator is then employed in four features selected above to make them between zero and one and then multiply. Finally, segmentation is performed according to the values obtained in step three and preset threshold. Result shows that the system not only achieves slightly better results than manual counting, but also provides reliable information of the bacteria shape. Figure 9 shows an example of segmentation result.
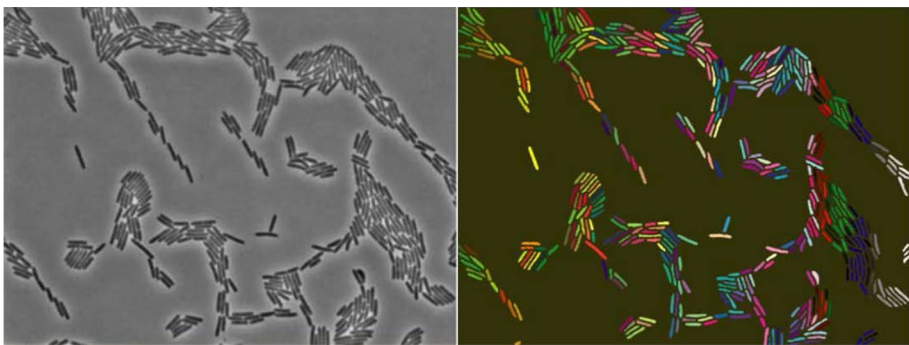

Fig. 9 The detection result of proposed method in Vallotton et al. (2010). The figure corresponds to Fig. 5 in the original paper

In Mukti et al. (2010), a method for detecting and counting Mycobacterium TB is proposed. Through this method, a threshold suitable for the image can be obtained. On this basis, the main goals of color threshold and image segmentation can be achieved. The first step is image segmentation by color threshold method. Counter is then used for counting out number of black Mycobacterium TB cells based on resulting binary image. By employing the method, the patient's disease grade determined by the number of TB cells can be accurately judged.

In Verikas et al. (2012), in order to protect the environment better, the system of detecting P. minimum automatically is proposed, which can get the quantitative concentration estimates of object. Firstly, an object image is created by segmenting Lab image. The center and contour of circular objects are then determined, respectively. After that, circular noise is eliminated based on the area threshold. For images with more than one object, the method takes a sequential detection method that removes each detection object. The developed algorithm is tested with 114 images recorded by a color camera. According to the test results, the introduced algorithms makes a good detection performance with a detection rate of 93.25%.

In Raof et al. (2011), for handling with large number of TB cases, a TB image segmentation method is proposed with the same accuracy as manual diagnosis and higher efficiently. The main idea includes two steps. One is image enhancement that involves adjusting image brightness, contrast, and color operations. The other is achieved by using color threshold technique. At this step, the color information is collected for TB pixels detection. An example of segmentation result is shown in Fig. 10. The final images suggest that TB in the original images can be detected accurately based on proposed method.

In Shi et al. (2012), to improve the efficiency of counting the total number of colonies, an automatic detection method of bacteria with high practical value is proposed. The main idea is to design a circular filter based on the features of low brightness around the live bacteria and high brightness in the center for detection. By comparing the results of manual detection and machine detection, it is shown that the recognition error of this system is within 10%.

In Rachna and Swamy (2013), an algorithm based on image processing is developed for identification of TB bacteria in sputum, which can reduce the workload of doctors in diagnosing tuberculosis. The first step is preprocessing. Then a global threshold is employed to remove tissue and background. After that, k-means clustering approach and
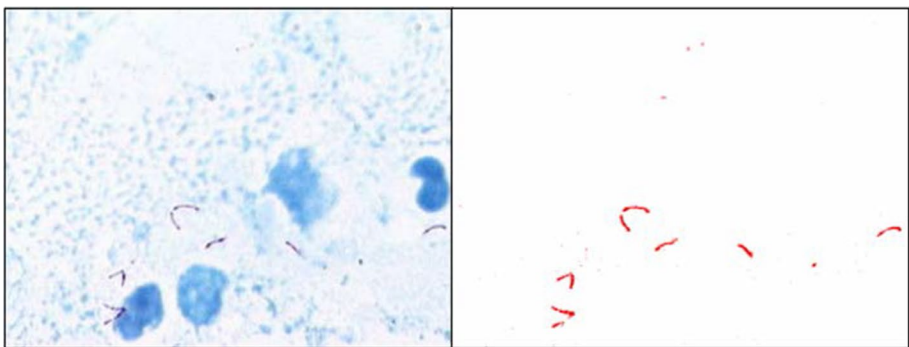

Fig. 10 The segmentation result of proposed method in Raof et al. (2011). The figure corresponds to Fig.6 in the original paper

Otsu threshold approach are respectively applied to detect TB for comparing their performances. At last, region growth is applied to mark and remove noise and large areas of excessive pollution after objects extracted by *k*-means clustering approach or Otsu threshold approach. The fully automatic TB bacilli segmentation method is tested on 25 positive tissue slides. The test result shows that the proposed method yields a satisfactory performance in TB segmentation with an accuracy of 98.00%.

In Matuszewski et al. (2013), for detection and tracking small objects in video streams, a set of spatial filters are designed. The first step is to calculate the Fourier transformation for the sample image. The second step is creating a binary filter by the values above the threshold parameter. Thirdly, the input image is transformed to the frequency domain and its spectrum is multiplied by the filter. After that, the inverse Fourier transformation of the result is calculated and a very blurred image is obtained. At last, other threshold value is applied to the output picture. Experiments show that the presented method fits for detecting and tracking known objects.

In Kowalski et al. (2014), an effective nematode extraction method is proposed for obtaining more informations about nematode behavior. First, the grayscale image is binarized using the adaptive threshold method. Second, a globle threshold is applied for extracting reference labels from original image. The XOR operation are then employed between the binary image obtained in first step and the binary image with extracted markers. At last, the position of nematode detected is determined by the center of the largest blob. Based on this method, the location of the nematode can be accurately determined. At the same time, with the aid of this method, it is possible to better track the nematodes.

In Wang et al. (2013), in order to replace manual observation with automatic observation based on machine vision, this paper proposes an edge detection method that is more suitable for low-contrast images. Canny operator is firstly employed for segmenting the low-contrast image while mathematical morphology is also applied in the backup image for the same motivation. Then two segmented images resulting from two method respectively are fused by employing wavelet transform. Experiment shows that the segmented image after fusion has better results than the previous image segmented based on a single method. It can be seen from Fig. 11 that the fused segmentation image can reflect the object feature more completely and clearly.

In Kurtulmuş and Ulu (2014), a new method using computer vision to detect and count the deaths of Heterobacter elegans from microscope images is proposed for better controlling agricultural pests and improving the detection efficiency of Heterobacter elegans. The first step is obtaining smooth medial axes of the nematode worms by a median filter, Otsu threshold, a morphological operation of area opening and filling operation. The overlapped nematode worms are then separated by a skeleton analysis. Finally, two different path analysis methods are used for detection of dead nematodes, which are both based on the standard error of the mean. A dataset containing total 685 images is prepared for verifying the performance of proposed methods, which includes 935 live worms and 780 dead one. The experiment result shows that the proposed algorithm can achieve a good performance with a detection rate of 85%.

In Farahi et al. (2015), in order to better diagnose visceral leishmaniasis, a method of segmenting leishman bodies is proposed. The first step is preprocessing by a linear contrast stretching transformation. In order to improve the performance of final segmentation, all small oval objects in image are extracted by applying morphological closing. In this step, objects, including leishman bodies and other shape-like objects are extracted. To extract boundary of objects detected, CV level set is employed. Finally, to remove
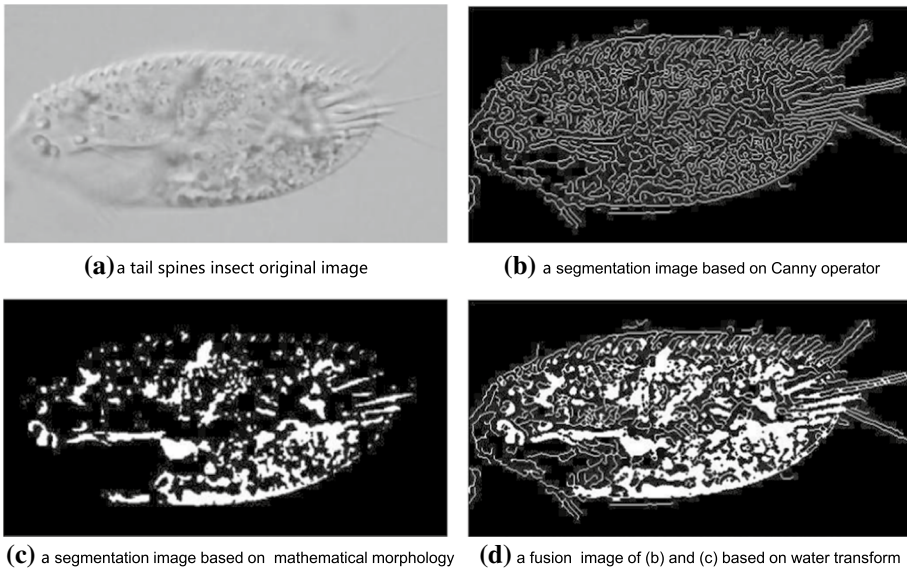
(a) a tail spines insect original image   **(b)** a segmentation image based on Canny operator

(c) a segmentation image based on mathematical morphology   **(d)** a fusion image of (b) and (c) based on water transform

Fig. 11 The segmentation results of proposed method in Wang et al. (2013). The figure corresponds to Figs. 1–4 in the original paper

the influence of other shape-like objects, a local threshold is applied. Experiment results suggest that the proposed method yields a mean of segmentation error of 9.76%.

In Goyal et al. (2015), for improving the efficiency of tuberculosis diagnosis, a reliable method is presented. The first step is converting the color space of input image into gray space. After that, in order to improve the performance of TB segmentation, tubeness filtering is employed for highlighting objects. The third step is roughly extracting bacteria based on Otsu global threshold. After labelling the extracted bacteria, all the other noise are removed. When compared to expert manual detection, the automatic detection algorithm presented in this work obtains a count of bacteria that shows a strong correlation with manual results.

In Shah et al. (2016), to improve the sensitivity and specificity of tuberculosis testing, an automatic detection method used on smart-phones is introduced. The first step is preprocessing that involves grayscale conversion and contrast enhancement. The second step is to binarize the image based on the threshold to separate the bacteria from the background. Morphological closing and filling are then performed. In step four, artifactas are removed based on geometric features. Finally, overlapping (touching) bacilli are segmented and separated by the watershed algorithm, after labelling TB. Total 30 images from smart-phone are prepared. One half is TB positive, the other half is TB negative. Experiment result shows that the presented method achieves a sensitivity of 93.3% and a specificity of 87%.

In Zhou and Liu (2016), a microbial contour extraction method is proposed to extract object contours from sewage microscopic images. First, the Sobel operator is used for to calculating the gradient of the image. Based on the calculated values, four adaptive thresholds are then employed for obtaining better object edges. After that, the extracted edges are connected based on the edge connection method designed. Finally, the outermost contour edge is then extracted from the image after processing of edge connection. Experiment

shows that compared with traditional edge detection algorithms, the proposed method can extract higher quality of microbial contours.

In Payasi and Patidar (2017), for improving the efficiency of tuberculosis diagnosis, a feasible algorithm is proposed. The first step is converting the image color space from RGB to HSI. The second step is segmenting the Hue component image. After denoising based on the area threshold, the holes are also filled. Finally, object is segmented from the processed image and the area and perimeter of object are calculated. Experiment results on prepared data show that the proposed method yields a high accuracy of over 90% in TB detection.

In Kemmler et al. (2011), for reducing the workload of labelling and counting bacteria in bright field microscopes, several semantic segmentation techniques are compared. One of introduced methods is region-based level set segmentation method. A database including five different microbe species with 40 up to 470 microbes per class is prepared for testing. The average recognition rate is employed for evaluating the detection performance. Other methods involved in this paper are explained in the corresponding chapters of their category.

3.2 Classification based methods

In this subsection, related works that use classification methods for microorganism detection are introduced. To present this works more clearly, we decide to introduce the involved methods separately according to their adopted features. Features commonly used in traditional image classification, as shown in Fig. 12, include shape features, geometric features, color features, texture features, statistical features, etc.

3.2.1 Classification based methods with shape features

Shape features are commonly used in traditional image detection methods, which mainly contain contour features and area features. More specifically, it include feature parameters such as squareness, angular, roundness, invariant moment, eccentricity, polygon description, and curve description.

In Dubuisson et al. (1994), Javidi et al. (2006), Yeom et al. (2006), and Liu et al. (2014a, b), contour feature, as a type of shape feature, is employed for microorganism detection. In
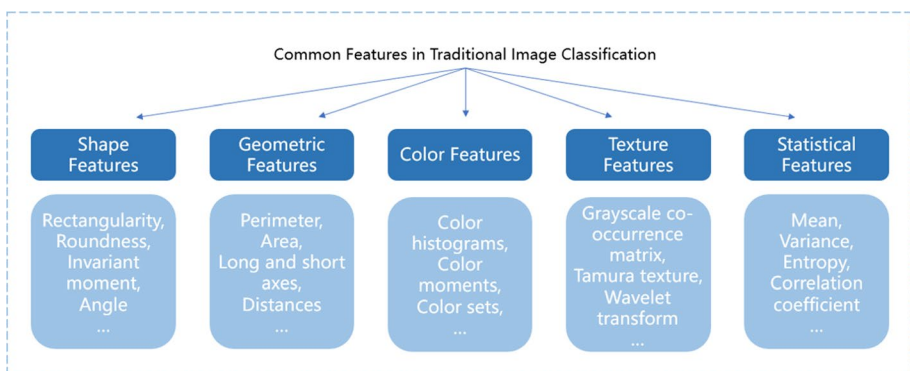

Fig. 12 Features commonly used in traditional image classification

Dubuisson et al. (1994), contour feature is extracted by Canny edge detector and threshold segmentation. Based on contour feature, Methanospirillum hungatei and Methanosarcina mazei are successfully distinguished and detected. Specific process of described algorithm is shown in Fig. 13. In Javidi et al. (2006) and Yeom et al. (2006), contour feature is combined with rigid graph matching (RGM) method to segment and detect biological microorganism in 3-D image. The segmentation results and detection results are respectively exhibited in Fig. 14. In Liu et al. (2014a, b), an optofluidic imaging system,which obtains contour feature and other informations by the image processing software called ImageJ, is proposed. In Liu et al. (2014a), the system is used for detecting *E. coli*, *Shigella flexneri* and *Vibrio cholera*. In Liu et al. (b), it is proposed to detect flu virus by the same research group with (Liu et al. 2014a). Experiments in Liu et al. (2014a, b) show that all the prepared microorganisms can be effectively detected and classified with proposed system.

In Rizvandi et al. (2008), Huang et al. (2008), Rizvandi et al. (2008a, b), and Zhou and Baek (2008), some similar methods based on angular features are proposed to detect *C. elegans* nematode worms. This methods contain several similar steps. The first step is preprocessing, including image binarization, small hole filling, skeleton and pruning. The second step is splitting skeleton into independent branches by processing specific pixels. Then, all the individual skeletons are reconstructed by applying branch merging method. At last, angular features extracted from proposed skeletons are used for *C. elegans* nematode worms detection. An outcome of experiment in Rizvandi et al. (2008a) is shown in Fig. 15. Related result shows that the best percentage of correctly detected worms is about 83%.

In Zhai et al. (2010), Hiremath et al. (2011), and Badsha et al. (2013), roundness feature is especially employed for detecting microorganisms that are approximately circular in shape. Method presented in Zhai et al. (2010) includes image segmentation and bacilli detection by roundness and roughness. Among them, roundness is used for distinguishing Single-bacillus objects and touching-bacillus objects. The experiment is carried out
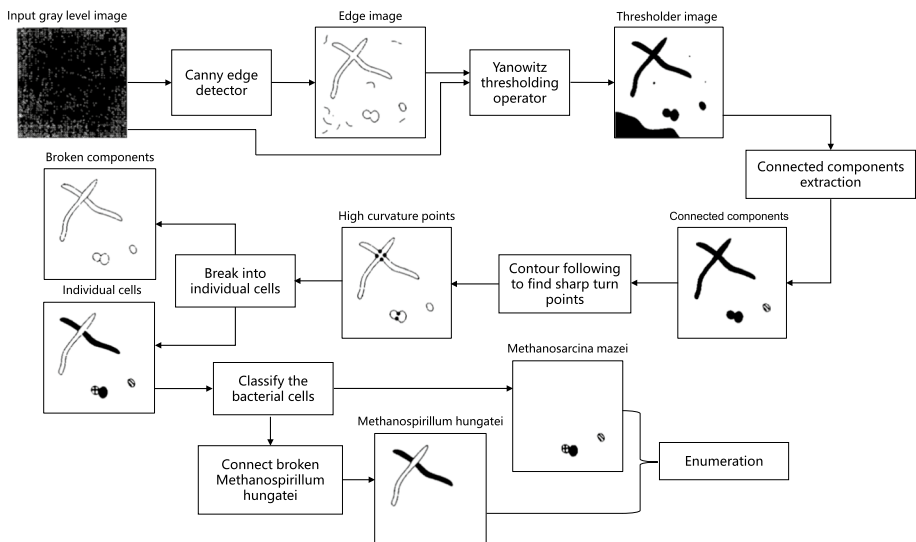

Fig. 13 The overview of proposed method mentioned in Dubuisson et al. (1994). The figure corresponds to Fig.1 in the original paper
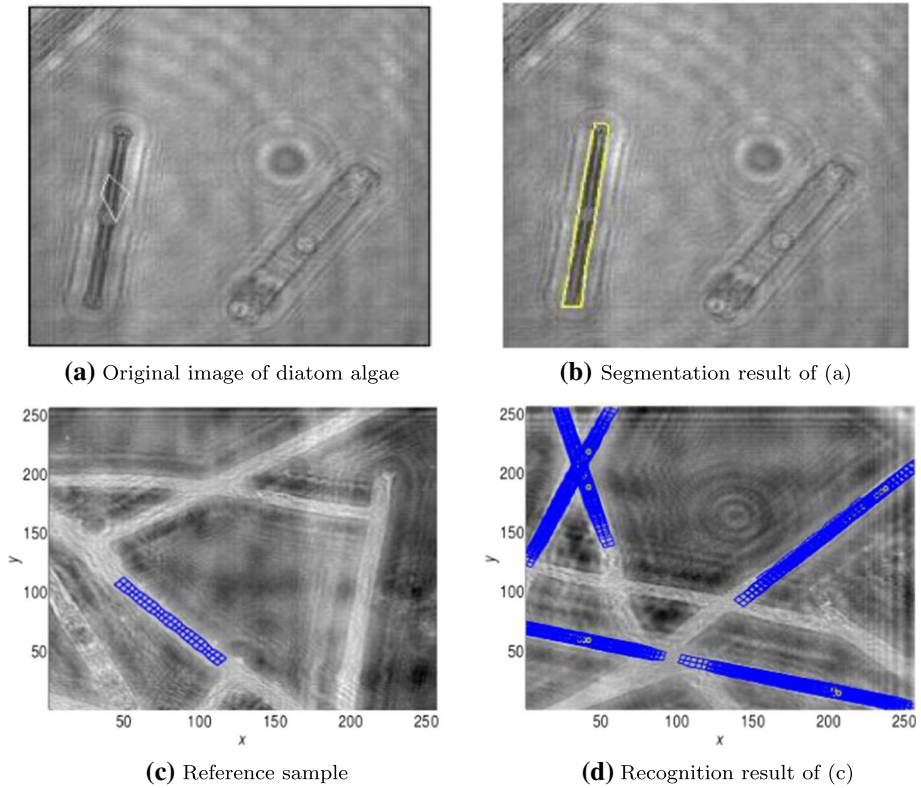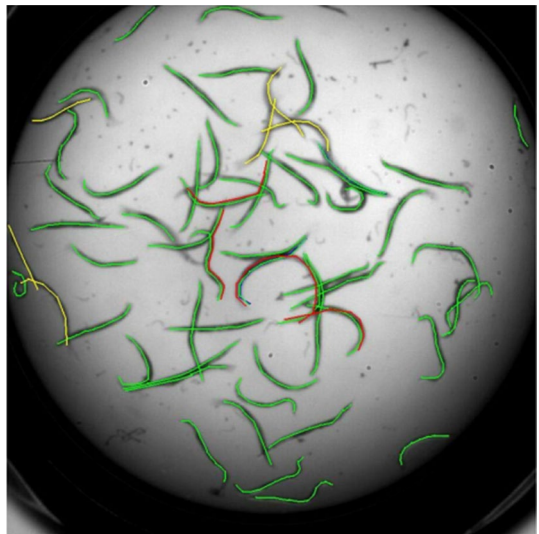
(a) Original image of diatom algae

(b) Segmentation result of (a)

(c) Reference sample

(d) Recognition result of (c)

Fig. 14 The detection results of proposed method in Javidi et al. (2006). The figure corresponds to Figs. 6 and 10 in the original paper

Fig. 15 The detection results of proposed method in Rizvandi et al. (2008a). The figure corresponds to Fig. 5 in the original paper

on 100 images. Result shows that 95% of total prepared samples yield a good detection accuracy of over 80%. In Hiremath et al. (2011), after several preprocessing operator, such as grayscale conversion, image intensity values adjusting and binarized, roundness data is obtained from binarized objects. At last, all Rotavirus-A particles are identified based on roundness and pre-set rule. 50 images are prepared for test the performance of presented algorithm. Experiment result shows that the identification rate of presented algorithm is 98%. In Badsha et al. (2013), roundness combined with eccentricity is applied for detecting *Cryptosporidium* and *Giardia (oo)cysts*. The corresponding detection process is shown in Fig. 16. 40 images containing *Cryptosporidium* and *Giardia (oo)cysts* are prepared for testing the performance of described method. Performance is evaluated by nucleus counting rates. In addition, in order to evaluate the performance of the method more objectively, its results are compared with manual counts. Experiment results indicate that proposed method achieves a detection rate of 97%.

3.2.2 Classification based methods with geometric features

Geometric features generally refer to the position and direction of object in image, as well as the perimeter, area, distance and other characteristics of objects. Geometric features are more intuitive and simple, and can play a good role in object detection.

In Fang et al. (2008), Liu et al. (2014a, b), and Yu et al. (2014), areas of objects are chosen as an important feature for accurate detection. In Fang et al. (2008), area and Neural Network (NN) are applied for automatically identifying TB in Acid-Fast stain sputum smears. Experiments on 44 Acid-Fast stain microscopic images with 533 TB show that the Perceptron has 100% sensitivity and 39.8% specificity; the feed-forward backpropagation has 97.8% sensitivity and 72.4% specificity. In Liu et al. (2014a, b), not only the shape features mentioned in the previous Sect. 3.2.1 is applied, but also area feature. Results indicate that method proposed in this research performs well in detecting a wide range of bacteria by combining shape features and area feature. In Yu et al. (2014), area is considered as an important indicator to detect the presence of bacteriophage and to estimate the
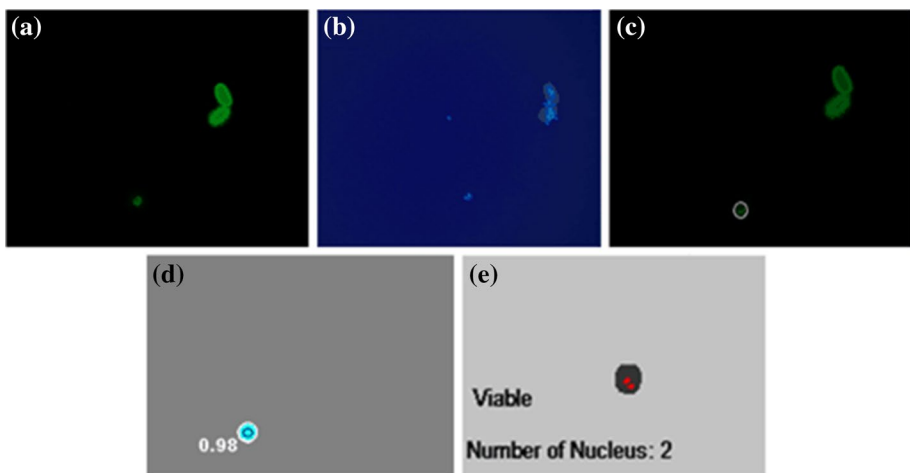

Fig. 16 Detection of *Cryptosporidium* and *Giardia (oo)cysts* in Badsha et al. (2013). The figure corresponds to Fig.8 in the original paper

number of bacteriophage. In Coltelli et al. (2013), centroid distance spectrum extracted from processed image is selected for microalgae detection. This methodology achieves a high accuracy of 96.6% in 3423 samples, which contain 24 kinds of microalgae. In Mäder et al. (2015), an image-analysis scheme using width information is presented to detect fungal infection. 415 images are prepared for evaluating the performance of proposed method with 194 positive samples and 221 negative one. Experimental results indicate that the proposed method yields a sensitivity of 83% and a specificity of 79% on prepared data.

In fact, the combination of several suitable geometric features selected for different microorganisms can achieve better detection. In Fukuda and Hasegawa (1989), Forero et al. (2003), Kumar and Mittal (2008), Coltelli et al. (2013), Mäder et al. (2015), Jan et al. (2015), Perner et al. (2004), and Sklarczyk et al. (2007), they achieve accurate and efficient detection of different microorganisms by using different combinations of geometric features. In Fukuda and Hasegawa (1989), a novel approach that considers a complete microbial individual as an individual composed of several basic shapes is proposed. Based on this basic shape's length and area features, the system matches the image with prepared data base. The identification experiment on three kind of microorganisms shows that the best detection performance is yielded in Nitzschia Fonticola with a detection rate of 90%. In Forero et al. (2003), many geometric features are randomly combined and tested for choosing the most suitable combination of features. At last, the combination of length, width and mahalanobis distance with a classification tree achieves the best result in bacilli detection. In Perner et al. (2004); Sklarczyk et al. (2007), the area-to-length ratio is applied for determining how close the target is to prepared template. The whole system architecture is shown in Fig. 17. Six different airborne fungi spores (shown in Fig. 18) are tested with proposed method. The highest recognition rate can be achieved for Scopulariopsis Brevicaulis with 98.2%. In Kumar and Mittal (2008), some geometric features of the region of interesting (ROI) are obtained by Image-Pro Plus for exploring a reliable and automatic microorganism detection technique, including width, length, area, perimeter and aspect ratio. In Jan et al. (2015), a method that applies area and bounding box is presented for greatly improving the efficiency of doctor diagnosis of tuberculosis. After color space conversion, several morphological operations and edge detection, data about area and
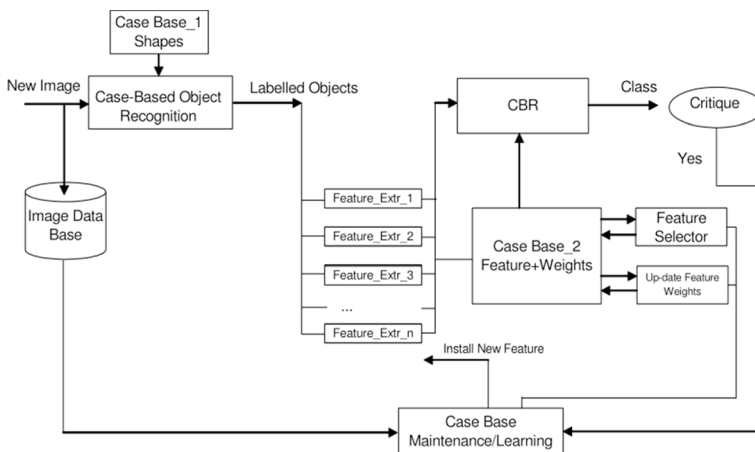

Fig. 17 The overview of proposed method mentioned in Sklarczyk et al. (2007). The figure corresponds to Fig. 6 in the original paper
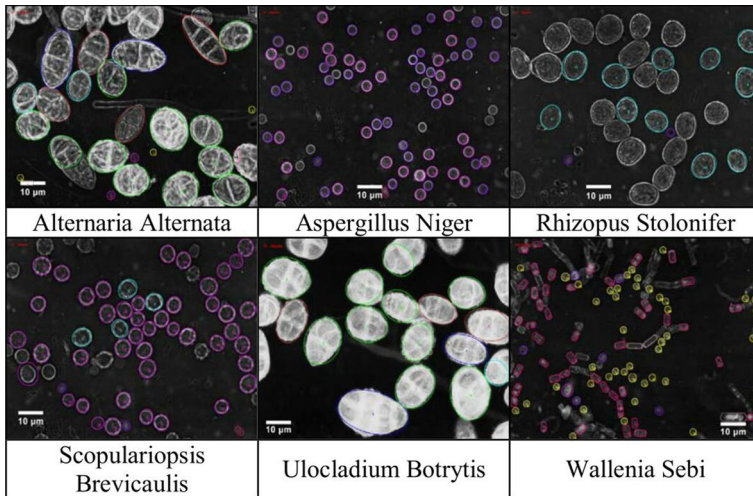
Fig. 18 The tested objects in Perner et al. (2004). The figure corresponds to Table 2 in the original paper

bounding box are obtained. Experiments on 100 positive samples and 10 negative samples show that it achieves a high accuracy rate of 90%.

3.2.3 Classification based methods with color, texture or statistical features

Color features can describe the surface properties of the corresponding scene in the image or image area. Color features description can be divided into color histogram, color distribution, color set, etc. In Ogawa et al. (2005), a novel multi-colour detection method is described for bacteria detection and metabolic activity assessment. Result shows that when compared to expert manual detection, the automatic detection algorithm presented in this work gains a count of bacteria that shows a strong correlation with manual results. In Tripathi et al. (2007), for detection and identification of individual bacterial cells and spores, a method based on Raman chemical imaging microscopy and color features is proposed. Result shows that Raman chemical imaging microscopy can distinguish a mixture of Gram-positive *Bacillus atrophaeus* spores and Gram-negative *E. coli* cells. In addition, color features are usually combined with other features for better detection performance. In Kumar and Mittal (2008) and Coltelli et al. (2013), color feature and geometric feature together constitute the feature parameters of the object. In Forero et al. (2003) and Fang et al. (2008), color features are regarded as the criteria for the initial screening of regions.

Texture features can reflect the slowly changing surface organization structure arrangement properties of the surface of objects. In Thiel and Wiltshire (1995), texture features are employed for detecting and distinguishing *Anabaena* and *Oscillatoria*. In this work, 14 *Anabaena cylindrica* and 20 *Oscillatoria agardhii* are prepared as the test group. The exoeriment result shows that the method distinguishes between *Anabaena* and *Oscillatoria* with an accuracy of over 90%. In Javidi et al. (2005), a real-time microorganism detection method by applying texture feature and RGM method is presented. Experiment result shows that the proposed method has strong performance on moving objects and different conditions. In addition, in Zhai et al. (2010), texture features are combined with shape features for TB detection.

Statistical features include mean, variance, energy, succession, etc. Statistical features are simple to compute and insensitive to the exact spatial distribution of color pixels. In Moon et al. (2010) and Javidi et al. (2010), a method based on statistical sampling methods is proposed for automated 3-D sensing and recognition of biological microorganisms. The specific steps are shown in Fig. 19. The test result shows that by comparing the variances, biological microorganisms can be accurately detected. In DaneshPanah et al. (2010), escape-force measurement Schaal et al. (2009) is combined with other cell identifications for detecting, classifying and controlling microorganisms and cells. In Shin et al. (2010), a system, which consists of a microfluidic device, a digital holographic microscope and relevant statistical recognition algorithms, is described for 3-D sensing and detection of microorganisms. Firstly, microorganisms are provided by the microfluidic device for the system to process. The Fresnel diffraction pattern of microorganisms is then optically recorded as a digital hologram. Thirdly, sampling segment features are randomly extracted from the reconstructed wavefront data. At last, statistical recognition algorithm is applied for cell identification. Microorganisms used in the experiments are *Euglena acus* and *Chilomanas*. Experimental results suggest that the optical fluid 3-D sensing and recognition method is feasible. In Yourassowsky and Dubois (2014), a system consisting of a Digital Holographic Microscope (DHM) and a partial coherent source is proposed for detecting and extracting objects in prepared samples. First, a non-zero complex amplitude is applied for determining whether the object exists. After that, the inverse Fourier transform and threshold operation are performed to obtain the location of the object. To assess the performance of DHM, tests are performed with a *Giardia lamblia* cysts image. Experiments show that the obtained phase and intensity images can be used for detection and classification.

3.2.4 Summary

Based on the above analysis of relevant studies on the use of classification methods for microorganism detection, we can find the that:

– The microorganism detection using classification methods has a development history of more than 30 years. Geometric features have been used as classification feature for microorganism detection in Fukuda and Hasegawa (1989) as early as 1989.
– Shape features and geometric features are widely chosen for microorganism detection due to their simplicity of calculation and ease of use.
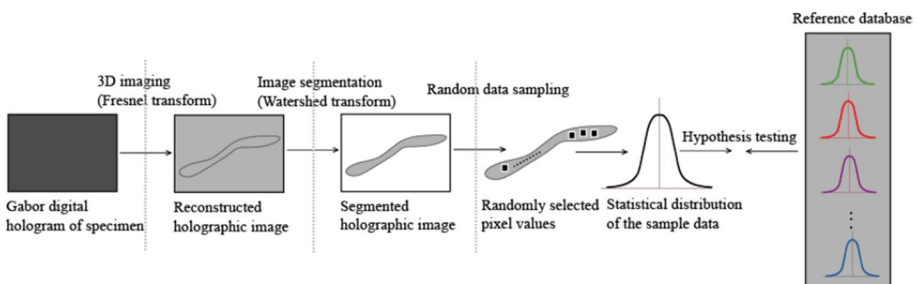

Fig. 19 The overview of proposed method mentioned in Moon et al. (2010). The figure corresponds to Fig. 2 in the original paper

– Using a combination of multiple features can achieve better detection results than a single feature.

3.3 Table analysis

In the past four decades, classical image processing methods for microorganism detection are continuously developed. All related works can be grouped into segmentation based methods and classification based methods. For better statistical related works, Table 4 is designed, where the publication date, literature links, research objects and its domains, main methods and evaluation indicators are presented.

4 Traditional machine learning based methods

Traditional machine learning, a popular research domain in recent decades, achieves good results in many computer vision tasks. Methods based on traditional machine learning are extensively tried for object detection. In this section, related works on microorganism detection based on traditional machine learning are introduced in chronological order in Sect. 4.1. After that, to show the progress of this researches more clearly in recent years, a concise table is prepared in Sect. 4.2.

4.1 Related works

In this subsection, related works based on traditional machine learning are introduced, including motivation, main methods, experimental data and results. In addition, some flow-charts are inserted to illustrate some of the research ideas better.

In Yin and Ding (2009), for detecting bacteria in vegetables, a recognition algorithm based on Back Propagation Neural Network (BPNN) is proposed. The first step is image segmentation based on iterative thresholds. The second step is to eliminate noise and extract object. Based on the extracted object, multiple morphological features of bacteria are obtained. At last, BPNN is applied for bacteria recognition by taking obtained feature data as input. The accuracy of presented algorithm is judged by comparison with the results of the conventional plate counting method. Experiment shows that in the test of 75 different vegetable samples, the correlation between the results of the two methods is 99.87%.

In Ochoa et al. (2010), an automatic identification method for *C. elegans* in population images is presented. A ridge segmentation method (Steger 1998) is applied for image segmentation in the first step. Shape and appearance data are then obtained by applying the open contour. At last, a probabilistic classifier is employed for *C. elegans* recognition. 687 *C. elegans* in 2000 linear objects are prepared for testing. Result shows that the best TP is 95%.

In Osman et al. (2010), a Genetic Algorithm-Neural Network (GA-NN) algorithm is proposed to assist pathologists in TB diagnosis. The first step is image segmentation, including color filter, moving *k*-means clustering and region growing. The second step is feature extraction based on Hu moment invariant technique and feature selection based on GA. The last step is classification using Neural Network (NN). 120 samples totally containing 360 TB and 600 possible TB are prepared. Among them, 200 TB and 200 possible TB are for training set. The others are equally grouped into the validation set and the test set. Result shows that the highest testing accuracy is 88.57%.

Table 4 Summary for object detection methods based on classical image processing

Date	References	Microorganism	Domain	Method	Evaluation
Segmentation based methods					
1985	Sieracki et al. (1985)	Bacterial	Water-borne	Object label	Compared with visual count
1988	Adams and Thomas (1988)	Hyphae	Medical	Segmentation, skeletonization	The overall percent Standard errors = 0.25%
1990	Packer and Thomas (1990)	Filamentous microorganisms	Medical	Binarization, circularity test, skeletonization	Compared with manual analysis
1991	Masuko et al. (1991)	Luminescent bacteria	Environment	Detection of objects that emit light	–
1995	Bloem et al. (1995)	Bacteria	Environment	Morphological operation, threshold segmentation	Compared with visual counts
1998	Baillieul and Scheunders (1998)	Microorganisms	Water-borne	Threshold	–
2003	Wang et al. (2003a, 2003b)	Cells from the first colonic gland of rats	Medical	Extracting nuclei, dividing image into regular non overlapped cubes	–
2006	Qing et al. (2006)	Algae	Environment	Threshold segmentation	Recognition rate ≥ 90%
2007	Li and Chen (2007)	Bacteria	Food	Histogram equalization, morphological operations, edge detection	–
2007	Coltelli et al. (2008)	Bacteria	Water-borne	Subtraction operation between two successive frames	–
2008	Costa et al. (2008)	TB	Medical	Color channel based difference operation, a global adaptive threshold	Sensitivity = 76.5%
2008	Zhang et al. (2008)	Bacterial colonies	Science	Dish/plate detection, Otsu	Satisfaction rate = 96%
Segmentation based methods					
2008	Wang et al. (2008)	Microorganisms	Water-borne	Multiple thresholds	–
2008	Fernandez et al. (2008)	Microorganism oocysts	Water-borne	Two threshold segmentation, color base detection	Correct detection rate = 100%
2009	Sotaquira et al. (2009)	TB	Medical	Threshold segmentation, a logical AND	Average efficiency = 96.3%

Table 4 (continued)

Date	References	Microorganism	Domain	Method	Evaluation
2010	Vallotton et al. (2010)	Bacteria	Science	Marr Hildreth edge detector, threshold segmentation	Compared with manual counts
2010	Mukti et al. (2010)	TB	Medical	Color thresholding segmentation	–
2011	Kemmler et al. (2011)	bacteria, fungi	Medical	Region-based level set	The average recognition Rate of yeast = 95%
2011	Raof et al. (2011)	TB	Medical	Image enhancement, Color thresholding segmentation	–
2012	Verikas et al. (2012)	P. minimum	Water-borne	Threshold segmentation, circular detection	Detection rate = 93.25%
2012	Shi et al. (2012)	Colonies	Food	Designing a circular filter	Recognition error ≤ 10%
2013	Matuszewski et al. (2013)	Microorganisms	Water-borne	Fourier transform, threshold, inverse Fourier transform	–
2013	Wang et al. (2013)	Protozoa	Environment	Canny edge detection, wavelet transform	–
2013	Rachna and Swamy (2013)	TB	Medical	Otsu thresholding approach	Accuracy = 98%
2014	Kowalski et al. (2014)	Nematode	Science	Adaptive thresholding	–
2014	Kurtulmuş and Ulu (2014)	Heterobacter elegans	Medical	Median filter, Otsu threshold, skeletonization	Recognition rates ≥ 85%.
Segmentation based methods					
2015	Farahi et al. (2015)	Leishman bodies	Medical	Morphological operation, CV level set	Mean of segmentation error for global model = 10.90%
2015	Goyal et al. (2015)	TB	Medical	Otsu global thresholding	Compared with expert manual detection
2016	Shah et al. (2016)	TB	Medical	Threshold segmentation, watershed algorithm	Sensitivity = 93.3%, Specificity = 87%
2016	Zhou and Liu (2016)	Microorganisms	Water-borne	Adaptive thresholds, edge connection	–

Table 4 (continued)

Date	References	Microorganism	Domain	Method	Evaluation
2017	Payasi and Patidar (2017)	TB	Medical	Color space conversion, segmentation, denoising	Counting accuracy = 90%
Classification based methods					
1989	Fukuda and Hasegawa (1989)	Microorganism	Industry	Basic shapes matching	Identification rate = 90%
1994	Dubuisson et al. (1994)	Microorganism		Pattern recognition	–
1995	Thiel and Wiltshire (1995)	Anabaena, Oscillatoria	Water-borne	Texture feature based classification	Accuracy ≥ 90%
2003	Forero et al. (2003)	TB	Medical	Canny operator, classification tree	Specificity = 96.09%
2004	Perner et al. (2004)	Fungi spores	Environment	Generating template, template matching	Recognition rate = 98.2%
2005	Javidi et al. (2005)	Alga	Water-borne	Threshold, RGM	–
2005	Ogawa et al. (2005)	Bacteria	Medical	Colour based classification	Compared with visual count
2006	Javidi et al. (2006); Yeom et al. (2006)	Sphacelaria alga, tribonema alga	Water-borne	RGM, statistical algorithms	–
Classification based methods					
2007	Sklarczyk et al. (2007)	Fungi	Air-borne	Case based matching	Recognition rate = 98.2%
2007	Tripathi et al. (2007)	Bacterial	Science	Raman chemical imaging microscopy	–
2008	Kumar and Mittal (2008)	Bacteria	Food	Analyzing geometrical parameters and three optical parameters	–
2008	Rizvandi et al. (2008a)	C. elegans	Science	Binarization, skeletonization	Correctly detected rate = 83%
2008	Rizvandi et al. (2008)	C. elegans	Science	Binarization, skeletonization	FRR = 7.9%, FAR = 8.4%
2008	Huang et al. (2008)	Nematode	Science	Adaptive threshold, morphological closing, skeletonization	–
2008	Rizvandi et al. (2008b)	Nematode	Science	Binarization, skeletonization	Detected accuracy = 89%

Table 4 (continued)

Date	References	Microorganism	Domain	Method	Evaluation
2008	Zhou and Baek (2008)	Nematode	Science	Binarization, skeletonization	–
2008	Fang et al. (2008)	TB	Medical	Color based detection, geometric features based classification	Sensitivity = 97.8%, Specificity = 72.4%
2010	Zhai et al. (2010)	TB	Medical	Combining two color space segment result	Detection accuracy ≥ 80%
2010	Moon et al. (2010); Javidi et al. (2010)	Microorganisms	Water-borne	Watershed transform, statistical sampling	–
2010	DaneshPanah et al. (2010)	Alga	Water-borne	Digital holographic microscopy	–
2010	Shin et al. (2010)	Protozoa	Water-borne	Microfluidic device, digital holographic microscope, statistical recognition algorithms	–
2011	Hiremath et al. (2011)	Rotavirus-A	Medical	Shape features based classification	Identification rate = 98%
Classification based methods					
2013	Badsha et al. (2013)	Protozoa	Water-borne	Threshold segmentation, roundness metric and eccentricity based classification	Identification rate = 97%
2013	Coltelli et al. (2013)	Microalgae	Water-borne	Centroid distance spectrum, dissimilarity measure and color based classification	Accuracy = 96.6%
2014	Yourassowsky and Dubois (2014)	Giardia lamblia cysts	Water-borne	High throughput filtering, Fourier transform, inverse Fourier transform	–
2014	Liu et al. (2014a)	Bacteria	Water-borne	Data of size, shape and refractive index based classification	–
2014	Yu et al. (2014)	Virus	Medical	An optical imaging technique	–

Table 4 (continued)

Date	References	Microorganism	Domain	Method	Evaluation
2015	Mäder et al. (2015)	Fungal	Medical	Statistical features based classification	The total sensitivity = 83%, Specificity =79%
2015	Jan et al. (2015)	TB	Medical	Threshold, edge detection,	accuracy = 90%

In White et al. (2010), a hierarchical approach is presented o measure the number of embryos, larvae and adults in an image. The proposed method is composed of four layers, which are respectively for finding the area of interesting, filtering and segmentation, breaking regions into object parts and object categorization, seen in Fig. 20. More than 1700 images of *C. elegans* are prepared. Precision and recall for the segmentation and labelling of each developmental stage are evaluated.

In Osman et al. (2010), for identification of TB in ZN-stained tissue image, an automatic method is described. The first step is initial filter to remove all the colour range except red. The second step is segmentation using moving *k*-mean clustering. In step three, moment invariant feature is extracted. The last step is detection by applying Hybrid Multi-layered Perceptron (HMLP) Network. 15 slides are prepared for experiment, each of which produces 30 to 50 images. Experiment results suggest that the proposed method achieve a accuracy of 98.07% and a specificity of 96.19%.

In Kumar and Mittal (2010), for identify pathogens in foods, an automatic and rapid detection method is proposed. The first step is background correction and isolating each cell of the treated image into the individual image. The second step is selecting the regions of interesting. After that, various geometrical, optical, and textural parameters of processed images are obtained. At last, the Probabilistic Neural Network (PNN) is designed and built for classifying the microorganisms. In the test on 155 images, PNN can classify the microorganisms using the nine best classification parameters with 100% accuracy.

In Khutlang et al. (2010), for TB detection, an automatic and reliable detection method is presented. For detecting possible bacillus, one-class pixel classifier is employed in stage one, where outputs images with color of object. Results of the mixture of Gaussian pixel classifier is shown in Fig. 21. The purpose of stage two is further segmenting image obtained in stage one. Eight samples containing 1064 positive objects and 1157 negative objects are prepared for test. Experiment result shows that the mixture of Gaussians classifiers performs best with the accuracy of 93.47%.
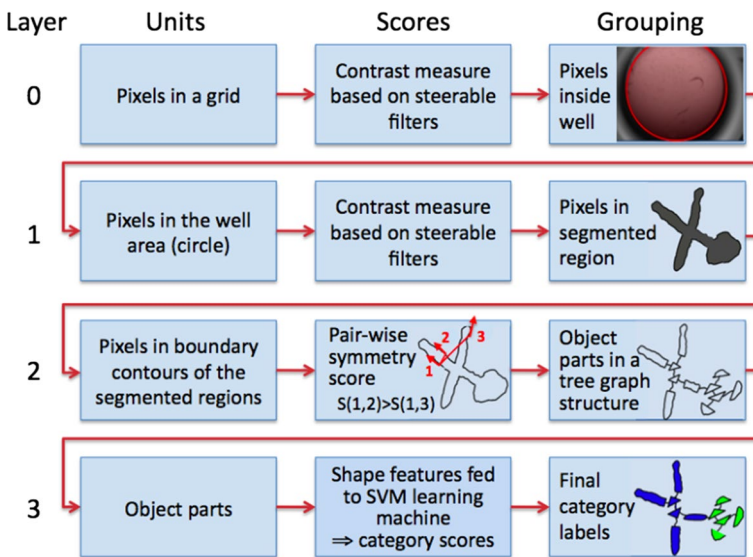

Fig. 20 The overview of proposed method mentioned in White et al. (2010). The figure corresponds to Fig. 2 in the original paper

Fig. 21 The detection results of proposed method in Khutlang et al. (2010). The figure corresponds to Fig. 1 in the original paper

In Zhang et al. (2010), an efficient detection method is proposed to automatically detect bacterial in foods. The first step is preprocessing, including subtraction between original image and background image, median filtering, gray-level histogram equalization. The second step is binarization on Otsu algorithm. The third step is filling the hollow area by morphological algorithm. At last, a support vector machine (SVM) classifier is applyed to classify the objects into the three categories. 50 images of the first category, 30 images of the second category and 20 images of the third category are prepared as training samples. Result shows that the relative error of results between proposed method and visual counts is less than 3%.

In Hiremath and Bannigidad (2010), to detection different kinds of cocci bacterial cells, an automatic method is proposed with low time-consuming and high accuracy. The first step is converting the original image into grayscale image and adjusting intensity values. The second step is image segmentation by active contour. The third step is labelling the segment image and computing geometric shape features for each labelled segment. The last step is applying 3σ classifier, k-Nearest Neighbor (k-NN) classifier and NN classifier to the feature set and outputting the classification of identified cells. 100 color images of each phase of bacilli bacterial cell are prepared for testing. Samples of test images are shown in Fig. 22. Result shows that the NN classifier yields 98–100% accuracy. In Hiremath and Bannigidad (2010), the same method is used by the same research group used for identifying and classifying the bacterial growth phases of bacilli cells. In addition, the Fuzzy classifier is also applied for this experiment. Result shows that the Fuzzy classifier yields 98–100% accuracy while the NN classifier yields 95–100% accuracy.

In Kemmler et al. (2011), which is already mentioned in Sect. 3.1, a segment-based classification using Conditional Random Fields (CRFs) is proposed. A database including five different microbe species with 40 up to 470 microbes per class is prepared for testing.

In Mansoor et al. (2011), for automatic detection four types of cyanobacteria genera of freshwater tropical Putrajaya Lake, an artificial neural network (ANN) based system is proposed. Steps of automatic algae recognition system are shown in Fig. 23. 400 samples of 4 cyanobacteria genera, namely *Microcystis*, *Oscillatoria*, *Anabaena* and *Chroococcus*, are prepared for experiment. Each genus has 100 samples that 80 samples for training and 20 samples for testing. The results illustrate as more than 95% success in identifying and classification the input samples of 4 genera of cyanobacteria.

In Osman et al. (2011a), for improving the detection performance and avoiding making false decision of computer-aided TB diagnosis system, an effective classifier is designed. The first step is image segmentation and k-means clustering algorithm based TB extraction. The second step is post-processing, including median filter, region growing and area
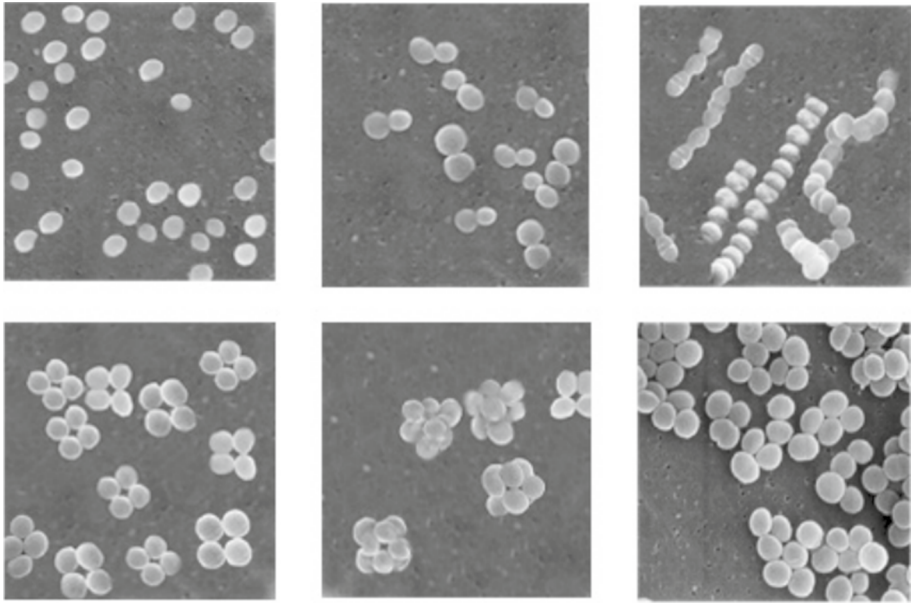
Fig. 22 The tested objects in Hiremath and Bannigidad (2010). The figure corresponds to Fig. 2 in the original paper
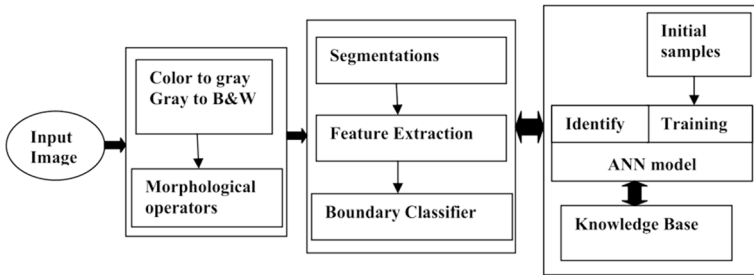

Fig. 23 The overview of proposed method mentioned in Mansoor et al. (2011). The figure corresponds to Fig. 2 in the original paper

based denoising. At last, a set of six affine moment invariants are calculated for every segmented region and are then fed into the Single Hidden Layer Feedforward Neural Network (SLFNN) for classifying the segmented regions into three classes: TB', 'overlapped TB' and 'non-TB'. 1603 objects combined of 'TB', 'overlapped TB' or 'non-TB' in 500 images are used for experiment (seen in Fig. 24), where 1000 objects are for training and 603 objects are for testing. Result shows that the SLFNN trained with the standard Extreme Learning Machine (ELM) method yields a high accuracy of 75.46%.

In Osman et al. (2011b), a SLFNN trained using the standard ELM algorithm is compared with Compact-SLFNN (C-SLFNN). Result shows that the classification accuracy of standard ELM is better than that of C-SLFN. However, standard ELM requires more hidden nodes than C-SLFN. In Osman et al. (2011c), the classification performance of Modified Recursive Prediction Error-ELM (MRPE-ELM) trained HMLP Network,

(a)

(b)

(c)
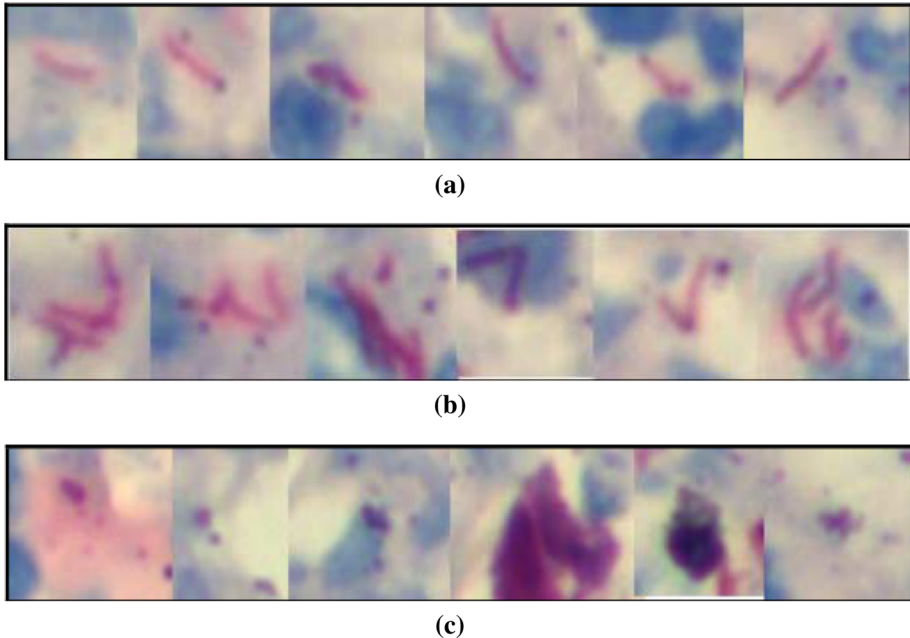
Fig. 24 The tested objects in Osman et al. (2011a). The figure corresponds to Fig. 3 in the original paper

MRPE trained HMLP Network and ELM trained SLFNN are compared together. Results suggest that MRPE-ELM has a better classification performance compared to MRPE algorithm.

In Hiremath and Bannigidad (2012), to detect different kinds of spiral bacteria, an automatic method is proposed with low time-consuming and high accuracy. The first step is converting the original image into grayscale image and adjusting intensity values. The second step is image segmentation by active contour. The third step is labelling the segment image and computing geometric shape features for all marked objects. The last step is applying 3σ classifier, k-NN classifier, NN and the Neuro Fuzzy classifier to the feature set and outputting the classification of identified cells. 300 color samples with three kinds of bacterial are prepares. Result shows that NN classifier as well as Neuro Fuzzy classifier yields a satisfied accuracy of 100% on prepared data.

In Ding et al. (2012), a rapid detection system with low detection costs is designed to meet the demand for rapid detection of *E. coli*. The first step is preprocessing, including creating a new memory area, grayscale the image, median filter denoising and threshold segmentation. Feature parameters are then extracted, including shape and color feature parameters. At last, Principal Component Neural Network (PCNN) is constructed and applied for *E. coli* detection. The results suggest that PCNN achieves a prediction accuracy of 91.33% on non-training samples.

In Mosleh et al. (2012), by combining image processing and ANN, an automatic approach for detecting some special freshwater algae genera is proposed. The first step is image preprocessing, including image resizing, image enhancement and noise removing. The second step is image segmentation based on Canny edge detection. In step three, shape and texture features are extracted and then normalized by applying principal component analysis. At last, Multilayered Perceptron (MLP) trained with back propagation error
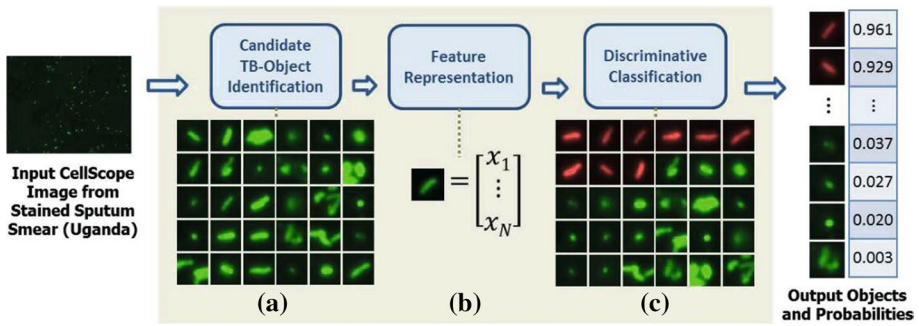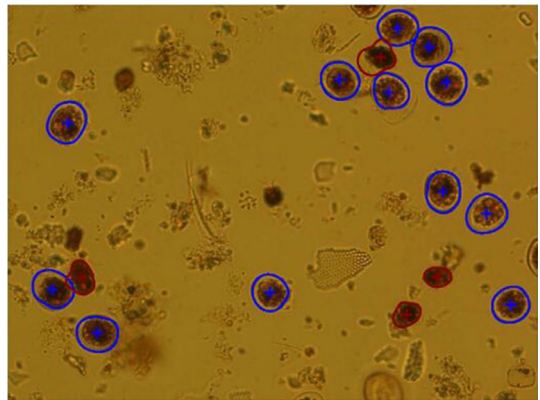
Fig. 25 The overview of proposed method mentioned in Chang et al. (2012). The figure corresponds to Fig. 3 in the original paper

Fig. 26 The detection result of proposed method in Verikas et al. (2012). The figure corresponds to Fig. 12 in the original paper

algorithm ANN is employed for classification. Experiment shows that the proposed method successfully detects 93 of 100 prepared samples.

In Chang et al. (2012), for detecting TB automatically, an algorithm with high accuracy is proposed, which is consisted of candidate TB identification step, feature representation step and discriminative classification step. Figure 25 shows the block diagram of designed algorithm. Experts mark TB objects in 92 of the 296 positive TB images, resulting in 1597 positive TB-objects. Each of the candidate objects is classified by an intersection kernel (IK)SVM, yielding an average precision of 91.3%.

In Verikas et al. (2012), a new technique is presented for detecting P. minimum species. In the first step, image edges is enhancement by applying phase congruency-based enhancement of image edges (Kovesi 2000). The second step is stochastic optimization-based object contour determination. At last, an automatic detection system combined of a Gaussian kernel SVM and a RF classifier is designed. Figure 26 exhibits the final detection result, where all P. minimum cells are detected. Experiment on 2088 P. minimum in 114 images shows that the proposed system achieves an detection accuracy of 93.2%.

In Zhai et al. (2012), to reduce the workload of doctors in diagnosing TB, a color and gradient feature based image segmentation and recognition algorithm is proposed. The algorithm mainly consists of two steps. One is image segmentation, including

pre-segmentation, adaptive segmentation and fusion segmentation. The other is image classification. Firstly, two gradient features and five shape features are extracted. Feature vector is then generated based on obtained features. At last, a Bayesian classifier is applied for classification. Experimental result shows that the object recognition rate of this algorithm can reach 91% on 100 images from different tuberculosis patients.

In Li et al. (2013), for recognition of environmental microorganisms (EMs), a framework of content-based image analysis is proposed with low cost and low time consumption. The first step is image segmentation. By comparing results of six segmentation approaches and considering the actual work needs, a semi-automatic segmentation approach employing Sobel edge detector is finally selected for the first step. The second step is shape feature description based on edge histograms, fourier descriptors, extended geometrical features, as well as internal structure histograms. At last, a multi-class SVM is applied for classification based on above features. The experiment is tested on a dataset with ten classes of EMs containing 20 images each. Ten samples from each class are for training, and the left 10 are for testing. Result shows that the best classification rate is 89.7%.

In Santiago-Mozos et al. (2013), an automatic screening systems is designed to detect agent in samples and make comprehensive decisions about subjects (e.g. ill/healthy) based on these tests. The system mainly consists of two stages. One is a classification stage. First, image is divided into many patches, which are then selected based on minimum green-color. Second, Canny edge detection is used for image segmentation. Third, a set of rotation and translation invariant features are extracted from each candidate object. At last, patches are classified by applying a SVM classifier. The other is a comprehensive decisions stage by applying a Bayesian methodology. The training set contains 34 TB-negative subjects and 11 TB-positive subjects while the test set contains 15 TB-negative subjects and 13 TB-positive subjects. Experiment result shows that the sensitivity of TB classifier is 73.53%.

In Xu et al. (2014), to improve the efficiency of doctor diagnosis, a computer image processing based algorithm of TB detection is proposed. The first step is preprocessing based on k-means algorithm. The second step is feature extraction. At last, Gaussian mixture model is trained with maximum expectation algorithm and then applied for classifying samples. The classification results of all data show that the classification sensitivity is 66.7%, and the accuracy is 96.2%.

In Zetsche et al. (2014), a system combined DHM and imaging-in-flow is described for the detection and classification of planktonic organisms. The first step is object localization from phase images, including classical threshold based XY-localization, robust refocusing criterion based Z-localization and re-segmentation based refinement of XY-coordinates. The second step is feature extraction containing texture features from both intensity images and phase images and morphological features from phase images. The last step is classification based on an SVM classifier. Results suggest that the correct prediction rate over all test data is 92.4%.

In Promdaen et al. (2014), a new method is proposed to detect 12 common microalgae in water resources of Thailand automatically. The first step is Sobel and Canny edge detection. Then, processed images are classified into rod-shaped and non-rod-shaped by sequential minimal optimization (SMO) classifier based on features of the axial ratio, the convex area and the area ratio. For rod-shaped algae image, the multi-resolution edge detection method is applied to re-segment. At last, SMO classifier is used for final classification. Experiment is performed on the data set that includes twelve genera of microalgae (seen in Fig. 27), 60 images for each genus (45 for training, 15 for testing). Experiment results suggest that proposed algorithm yields an accuracy of 97.22%.
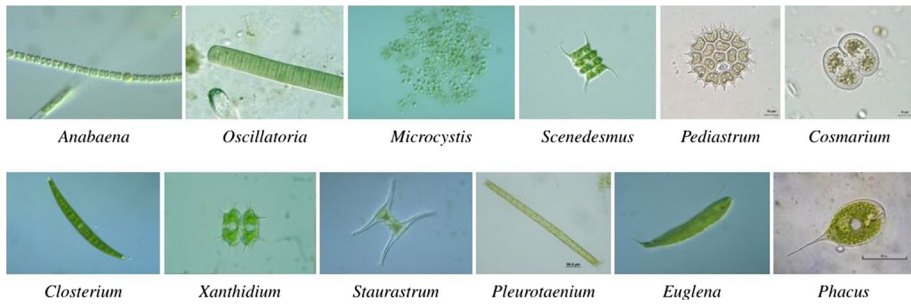
| Anabaena | Oscillatoria | Microcystis | Scenedesmus | Pediastrum | Cosmarium |

| Closterium | Xanthidium | Staurastrum | Pleurotaenium | Euglena | Phacus |

Fig. 27 The tested objects in Promdaen et al. (2014). The figure corresponds to Fig. 1 in the original paper

In Yang et al. (2014), for recognition of EMs, a framework of content-based image analysis is proposed with low cost and low time consumption. In addition, a novel 2-D feature descriptor is especially introduced for EM shapes. The first step is image segmentation by employing a Sobel edge detector. In the second step, the designed shape descriptor is applied for feature extraction. At last, a multi-class SVM is used for classification. The experiment is tested on a dataset with ten classes of EMs containing 20 images each. Ten samples from each class are for training, and the left 10 are for testing. Based on proposed feature descriptor, the overall classification rate increases by 2.8% from 89.7%.

In Verikas et al. (2014), an integrated approach is proposed for P. minimum detection. Firstly, histogram-based image binarization technique is applied to process the original images. Then, features for classification are determined, categorized and computed. At last, a committee combined of SVM and RF is applied to make a decision. The proposed method is performed on 158 images with 920 P. minimum cells. Result shows that the proposed method yields an overall recognition rate of 97% for P. minimum cells.

In Nugroho et al. (2015), for detection of malaria parasite cell, a method based on image processing is proposed. The first step is image enhancement, which contains contrast stretching and median filter. The second step is image segmentation based on k-means algorithm. Histogram-based texture is then extracted for the last step, in where multilayer perceptron backpropagation algorithm is applied for final classification. The prepared data set contains of 60 images which are grouped into trophozoite, schizont and gametocyte. The results suggest that this algorithm achieves a good detection performance in the prepared data with accuracy of 87.8% and specificity of 90.8%.

In Li et al. (2015), an EMs classification method is described, which solves the problem of small training data sets and noisy images. The presented algorithm contains three steps as following: sparse coding features extraction, region-based (RB)SVM designing and final localization and classification. Figure 28 shows the comparison of basic and improved RBSVM. The database contains 15 classes of microbes, each with 20 images (10 for training, 10 for testing). Mean of average precisions (MAP) is chosen as the evaluation measure. Result shows that the MAP of (RBSVM + NNSC) is also higher than that of (RBSVM + BoVW).

In Shan e Ahmed Razaa et al. (2015), an anisotropic tubular filtering (ATF) based algorithm is proposed for automatic detection of TB. Firstly, ATF is employed for image enhancement, which is mainly increasing the contrast between objects and background. The possible acid-fast bacillis (AFBs) are then segmented based on color threshold. Fourier descriptors and Hu's moment of each candidate AFB are then calculated. Finally, a region-based function (RBF)SVM is employed to classify each candidate AFB
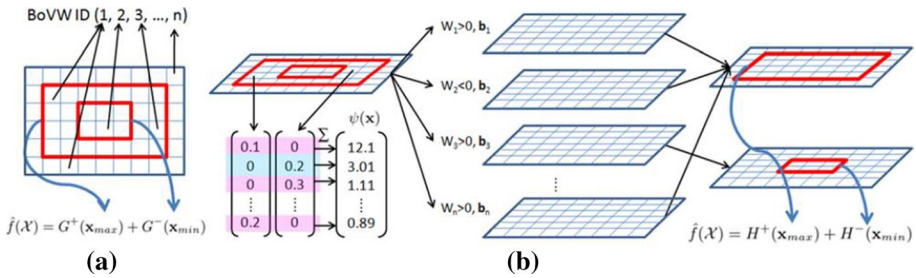
Fig. 28 The architecture of RBSVM mentioned in Li et al. (2015). The figure corresponds to Fig. 2 in the original paper

according to calculated features. Among 300 samples, 161 are marked as positive and the other 139 are marked as negative by experts. In addition, 180 images are grouped into training set while others are grouped into testing set. Comparative result shows that proposed method with Fourier descriptor has the highest sensitivity, F1-Score and accuracy.

In Dannemiller et al. (2015), a new method based on Retinex and SVM is presented for the segmentation of alga images. The proposed algorithm mainly contains two steps. One is improving image quality by applying the Retinex filtering technique. The other is segmenting algae in the image based on SVM. In addition, 100 samples containing only algae and 100 samples containing only background are obtained from prepared 32 images. Result shows that presented algorithm yields an overall detection rate of over 95%.

In Sajedi et al. (2019), for detecting actinobacterial species, an algorithms based on principle component analysis (PCA) and MLP is presented. The first step is data augmentation by applying Gaussian blurring. In the second step, feature vector is generated, including wavelet transform and dimension reduction. At last, feature vectors obtained above are classified by MLP. The experiment is performed on the database, named UTMC.V2.DB. In addition, classes with less than eight images are not considered. Results suggest that the accuracy is about 80.5%.

In Dhindsa et al. (2020), for identification of algae in water bodies, some classification algorithms are used, separately. Firstly, the gray scale images are subjected to the pixel clustering. Boundaries of the microbes are then extracted by using Otsu method and Kirsch filter. Thirdly, the object features corresponding to different classifiers are extracted. At last, the following classifiers are used for final classification: comparing Classification and Regression Trees (CART), k-NN, Gaussian Naive Bayes, Linear Regression, Linear discriminant analysis and SVM. The experiment to test the performance of the classifiers is carried out on 10 kinds of algae. Experimental evaluation and classification algorithm research show that the CART algorithm is the most suitable.

4.2 Summary

In recent decade, detection methods based on traditional machine learning are increasingly used for microorganism detection. The total related works are shown in Table 5, where the publication date, literature links, research objects, main methods and evaluation indicators are displayed.

Table 5 Summary for object detection method based on traditional machine learning

Date	References	Microorganism	Method	Evaluation
2009	Yin and Ding (2009)	Bacteria	Iterative thresholds, BPNN	Correlation = 0.99.87%
2010	Ochoa et al. (2010)	*C. elegans*	Open contour, aprobabilistic classifier	TP = 95%
2010	Osman et al. (2010)	TB	*k*-means clustering, GA-NN	Accuracy = 88.57%
2010	White et al. (2010)	*C. elegans*	A hierarchical approach	Precision vs. recall plots
2010	Osman et al. (2010)	TB	Initial filter, moving *k*-means clustering, HLMP Network	Accuracy =98.07%, Sensitivity = 100%, Specificity = 96.19%
2010	Kumar and Mittal (2010)	Five microorganisms	Background correction, PNN	Accuracy=100%
2010	Zhang et al. (2010)	Bacterial	Otsu algorithm, SVM	Relative error ≤ 3%
2010	Khutlang et al. (2010)	TB	Mixture of Gaussian classifier	Accuracy = 93.47%
2010	Hiremath and Bannigidad (2010)	Cocci bacterials	Active contour, 3σ classifier, *k*-NN classifier and NN	Accuracy of NN classifier = 98% to 100%
2010	Hiremath and Bannigidad (2010)	Bacilli cells	Active contour, 3σ classifier, *k*-NN classifier, NN and Fuzzy classifier	Accuracy of fuzzy classifier = 98% to 100%
2011	Mansoor et al. (2011)	Four types of cyanobacteria	ANN model	Classification rate = 95%
2011	Osman et al. (2011a, b, c)	TB	*k*-means clustering,region growing, SLFNN, ELM, C-SLFNN, MRPE, HMLP	Accuracy
2011	Kemmler et al. (2011)	Microorganisms	CRF	Average recognition rate of *E.coli* = 85%

Table 5 (continued)

Date	References	Microorganism	Method	Evaluation
2012	Hiremath and Bannigidad (2012)	Different spiral bacterial cells	Active contour, 3σ classifier, k-NN classifier, NN and Neuro Fuzzy classifier	Accuracy of NN classifier = 100%, Accuracy of Neuro Fuzzy classifier = 100%
2012	Ding et al. (2012)	*E. coli*	Median filter denoising, threshold, PCNN	Prediction accuracy = 91.33%
2012	Chang et al. (2012)	TB	Template matching, IKSVM	Average precision = 91.3%
2012	Mosleh et al. (2012)	Freshwater algae genera	Image enhancement, Canny edge detection, MLP, ANN	Accuracy = 93%
2012	Verikas et al. (2012)	P. minimum	Edge enhancement, SVM, RF	Accuracy = 93.2%
2012	Zhai et al. (2012)	TB	Adaptive segmentation, fusion segmentation, Bayesian classifier	Recognition rate = 91%
2013	Li et al. (2013)	EMs	Sobel edge detector, multi-class SVM	Classification rate = 89.7%
2013	Santiago-Mozos et al. (2013)	TB	Threshold, Canny edge detection, Bayesian methodology	Sensitivity = 73.53%.
2014	Xu et al. (2014)	TB	k-means, Gaussian mixture Model	Sensitivity = 66.7%, Accuracy = 96.2%
2014	Zetsche et al. (2014)	Planktonic organisms	Robust Refocusing Criterion, SVM	Prediction rate = 92.4%
2014	Yang et al. (2014)	EMs	Sobel edge detector, a new feature descriptor, multi-class SVM	Classification rate = 92.5%

Table 5 (continued)

Date	References	Microorganism	Method	Evaluation
2014	Promdaen et al. (2014)	12 common microalgae	Sobel edge detection, Canny edge detection, SMO classifier	Accuracy = 97.22%
2014	Verikas et al. (2014)	P. minimum	Histogram-based binarization, SVM, RF	Overall recognition rate = 97%
2015	Nugroho et al. (2015)	Malaria parasite cell	k-means, MLP	Accuracy = 87.8%, Sensitivity = 81.7%, Specificity = 90.8%.
2015	Li et al. (2015)	EMs	RBSVM, sparse coding features	MAP
2015	Shan e Ahmed Razaa et al. (2015)	TB	ATF, RBFSVM	Sensitivity, Specificity, F1-Score, Accuracy
2015	Dannemiller et al. (2015)	Algae	Retinex filtering technique, SVM	Overall detection rate ≥ 95%
2019	Sajedi et al. (2019)	Actinobacterial strains	PCA, MLP	Accuracy = 80.5%.
2020	Dhindsa et al. (2020)	Algae	Otsu, Kirsch filter, multiple machine learning classifiers	Accuracy, Recall, Precision, F1-score

5 Deep learning based methods

Recently, deep learning develops rapidly and achieves a series of excellent results in image analysis. In this section, we first introduce the advantages of deep learning methods. Then, works related to microorganism detection based on deep learning are summarized. After that, a concise table is designed to show the progress of deep learning in microorganism detection more clearly.

5.1 Advantages of deep learning methods

Compared with traditional machine learning methods, deep learning methods have a wide range of applications and strong applicability. Deep learning methods build its network framework by combining multiple simple but nonlinear modules, which enables it to design corresponding module combinations to achieve function mapping according to different problems LeCun et al. (2015). In addition, In the feature extraction step of detection processing, traditional machine learning methods apply handcrafted feature engineering methods, which is labor-intensive and time-consuming. Deep learning methods cannot only achieve automatic learning of features through its advanced network structure, but also learn complex features from simple features. Next, the advantages of deep learning for object detection are better represented by introducing some backbones of deep learning networks.

In Simonyan and Zisserman (2014), the VGG backbone is proposed, which has two structures: VGG-16 and VGG-19. The deepening of network layers enables VGG to obtain larger feature maps. In addition, the superposition of multiple small convolution kernels enhances the feature learning ability of VGG. Relevant comparative experiments show that VGG can learn more complex and more feature information than the previous traditional machine learning. VGG is the backbone of single shot detector (SSD) Liu et al. (2016), a object detection model with high speed and high accuracy. In Szegedy et al. (2015), Goog-LeNet is presented, which is based on Inception modules shown in Fig. 29. The feasibility of obtaining optimal sparse structure with ready-made dense building blocks is confirmed by GoogLeNet. GoogLeNet not only deepens the depth and width of the network, but also reduces the amount of parameters. In He et al. (2016), a residual learning framework called
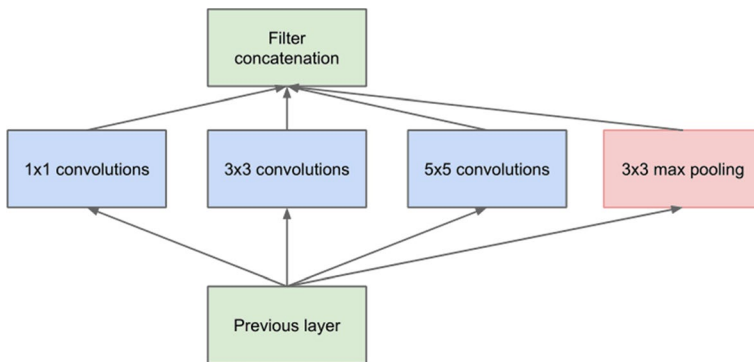

Fig. 29 Inception module mentioned in Szegedy et al. (2015). The figure corresponds to Fig. 2 in the original paper

Fig. 30 Residual block mentioned in He et al. (2016). The figure corresponds to Fig. 2 in the original paper
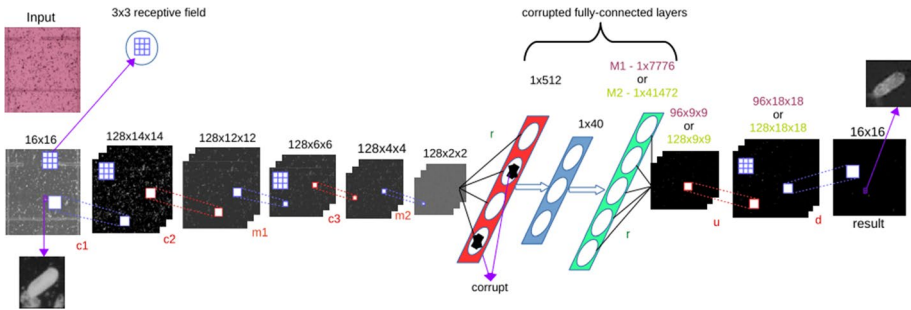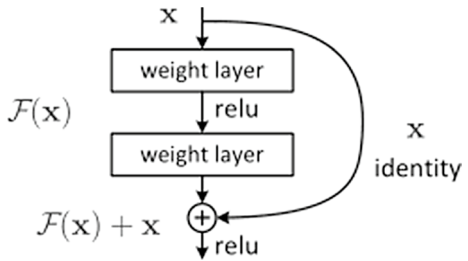

Fig. 31 The architecture of proposed model mentioned in Akintayo et al. (2016). The figure corresponds to Fig. 3 in the original paper

ResNet is proposed to solve the problem that training difficulty will increase with the deepening of network layer. By employing residual block, shown in Fig. 30, the problem of gradient disappearance that occurs when training deeper network layers is well solved. With the backbone of ResNet, both Faster R-CNN Ren et al. (2015) and Mask R-CNN He et al. (2017) achieve great detection results.

5.2 Related works

In this subsection, related works based on deep learning are introduced, including motivation, main methods, experimental data and results. In addition, some flowcharts are inserted to illustrate some of the research ideas better.

In Akintayo et al. (2016), a new selective autoencoder algorithm based on deep convolutional networks is designed to detect soybean cyst nematode eggs. Firstly, the proposed deep convolutional network is trained with training patches, where error back-propagation algorithm is applied. secondly, the trained network is tested with test patches and then outputs test results. Convolutional autovaluational architecture for two alternative model structures are shown in Fig. 31. The bounding box of every soybean cyst nematode egg in 644 images are extracted and stored for experiment, of which 80% are used for training and 20% are used for verification. Result shows that the average detection accuracy of proposed approach is 94.33%.

In Hung and Carpenter (2017), a faster region-convoluational neural network (Faster R-CNN) is firstly employed for detecting cell in blood and determining their stages. Firstly, a traditional segmentation method and machine learning are employed for this task. Based on their performance, the baseline is established. After that, a two-stage
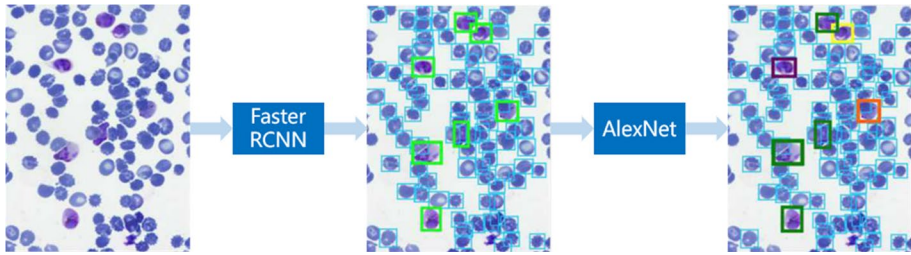
Fig. 32 The overview and detection result of proposed method mentioned in Hung and Carpenter (2017). The figure corresponds to Fig. 4 in the original paper
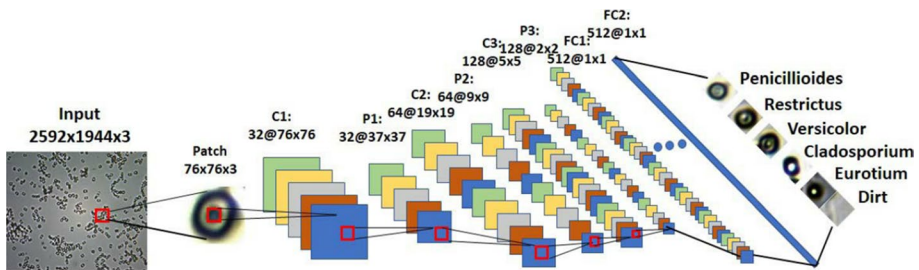

Fig. 33 The architecture of CNN mentioned in Tahir et al. (2018). The figure corresponds to Fig. 5 in the original paper

detection and classification approach is selected, seen in Fig. 32. In stage one, Faster R-CNN is applied for classifying bounding boxes around objects as red blood cell (RBC) or non-RBC. In stage two, non-RBCs detected in stage one are sent to AlexNet for further classification. 1300 images containing around 100,000 single labelled cell is prepared. Experiment results suggest that the two-stage algorithm yields a total accuracy of 98%.

In Tahir et al. (2018), a CNN-based approach for detecting fungus and distinguishing different types of fungus is proposed. In addition, a novel fungus dataset consisting of five different types of fungus spores and dirt is developed. The proposed CNN architecture for detecting fungus is summarized in Fig. 33. The presented method is trained with 30,000 images (5000 images for each class) and test on 10,800 images (1800 images for each class). Result shows that the accuracy of proposed method is 94.8%.

In Panicker et al. (2018), for detecting TB in microscopic sputum smear images, a CNN-based detection method is presented. The proposed method mainly consists of two stages. One is an image binarization stage by using Otsu threshold algorithm. The other is a pixel classification stage by applying CNN to determine the class of regions extracted from stage one. 22 microscopic images containing high-density and low density images are prepared. Segmentation and classification results of TB shown in Fig. 34 suggest that the presented method yields a recall of 97.13%, a precision of 78.4% and a F-score of 86.76%.

In Treebupachatsakul et al. (2016), a CNN based system for bacteria detection is proposed to reduce the analysis time and increase the accuracy of diagnostic process. 400 sample images of *Staphylococcus aureus* (*S. aureus*) and 400 sample images of *Lactobacillus delbrueckii* (*L. delbruekii*) are collected. Each type images are separated into 20% training datasets and 80% test datasets. Result shows that the best validation accuracy is 96%.
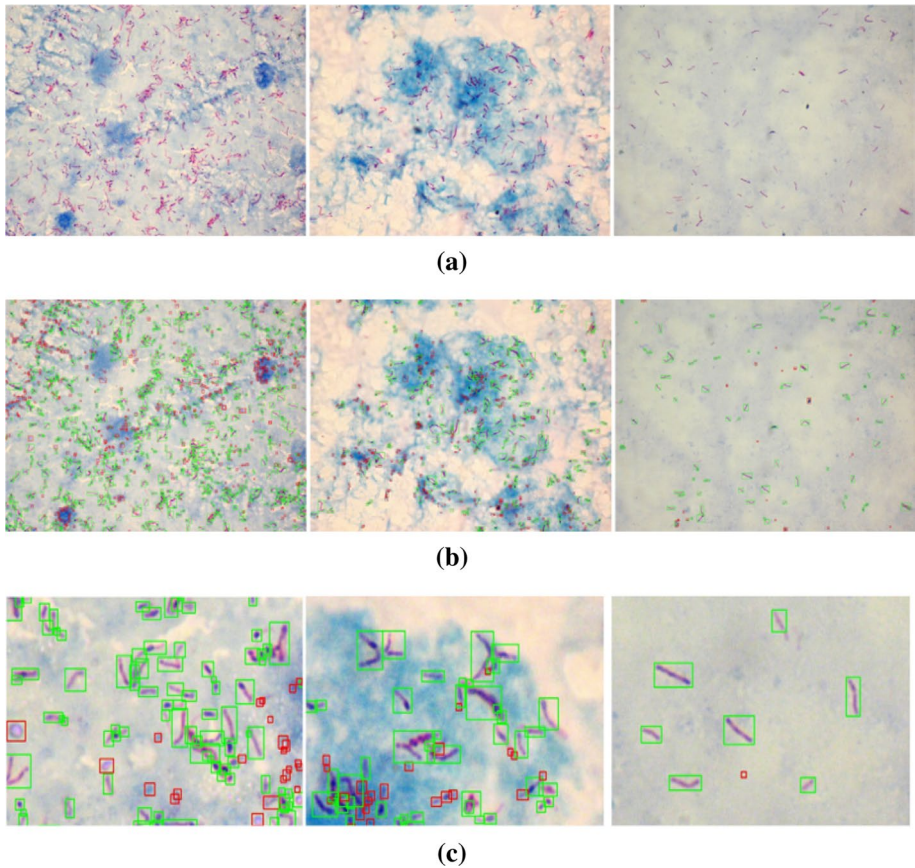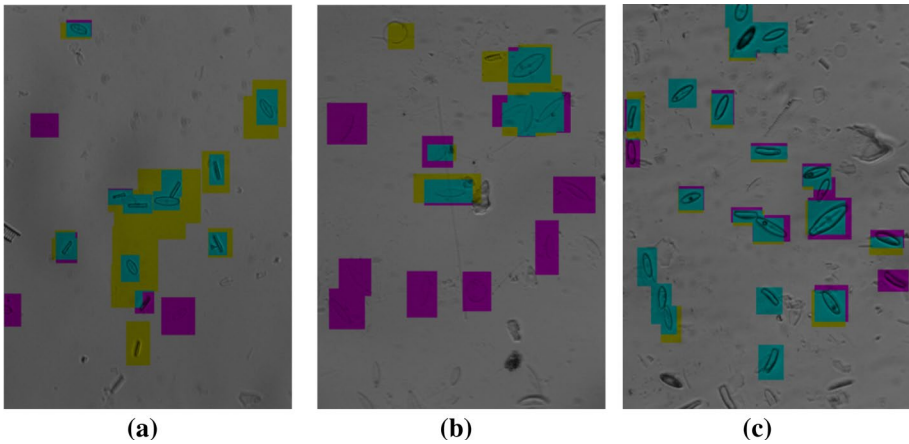
Fig. 34 The detection results of proposed method in Panicker et al. (2018). The figure corresponds to Fig. 5 in the original paper

In Pedraza et al. (2018), the performance of R-CNN and you only look once (YOLO) in diatom detection is compared. The R-CNN based method consists of four fundamental steps as following, generating edge boxes for region proposal, rejecting proposed regions, classification of regions and box merging. The YOLO based detection method also contains four steps. Firstly, the input image is divided into a cell matrix and each cell proposes a fixed number of candidate regions. The second step is moving each box to fit to a candidate object. After that, each box is classified and provided with a confidence score. At last, most of boxes are rejected using a threshold. The two methods are test on 11000 images from the 10 species. Detection results of RCNN and YOLO are shown in Fig. 35. Result shows that YOLO has better performance in diatom detection with a top F-measure of 84% compared to R-CNN.

In Sajedi et al. (2019), two deep learning based method is proposed to detect actinobacterial species on solid culture plates. In the first method, CNN is applied for actinobacterial strains detection. The second method consists of two steps. One is a data augmentation step, including blurring, cropping, and horizontal rotation. The other is a classification step based on ResNet, which training mechanism is transfer learning. Two

(a)   **(b)**   **(c)**

Fig. 35 The detection results of proposed method in Pedraza et al. (2018). The figure corresponds to Fig. 6 in the original paper

dataset respectively named UTMC.V1.DB and UTMC.V2.DB are prepared. UTMC.V1.DB contains 703 images from 55 different classes. UTMC.V2.DB contains 1303 images from 97 different classes. Result shows that the first method achieves up to 80.81% and 84.81% accuracy on UTMC.V1.DB and UTMC.V1.DB, respectively. The second method yields a accuracy of 90.24% in UTMC.V1.DB and a accuracy of 85.96% in UTMC.V2.DB.

In Viet et al. (2019), to detect parasite eggs, an automatic algorithm by applying Faster R-CNN is proposed. Figure 36 displays the architecture of Faster R-CNN. In addition, the decision metric of bounding box regression is that the intersection over
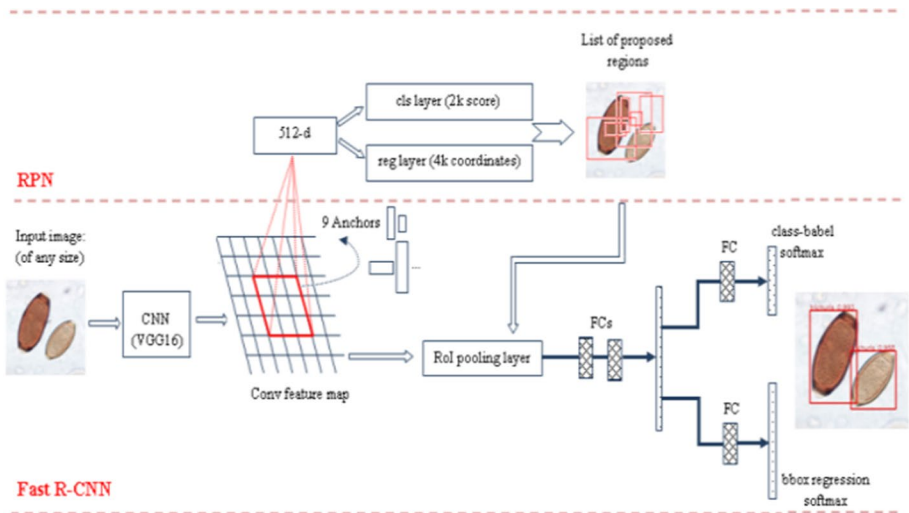
Fig. 36 The architecture of Faster R-CNN mentioned in Viet et al. (2019). The figure corresponds to Fig. 2 in the original paper
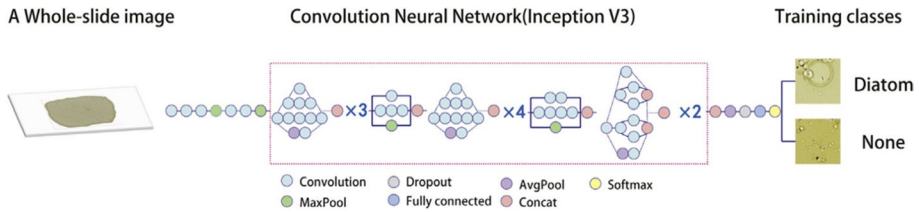
Fig. 37 The architecture of Inception V3 mentioned in Zhou et al. (2019). The figure corresponds to Fig. 1 in the original paper
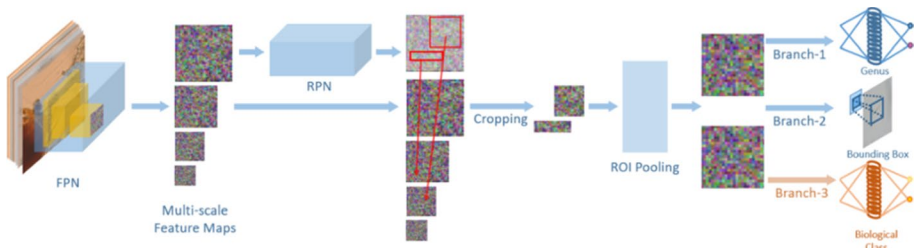

Fig. 38 The architecture of model mentioned in Qian et al. (2020). The figure corresponds to Fig.2 in the original paper

union must be greater than 70%. Images obtained from stool samples of patients are prepared for testing. Result shows that the mAP of Faster R-CNN based detection is very high with 97.67%.

In Zhou et al. (2019), an automatic detection system is proposed for diatoms classification. For the proposed method, a GoogLeNet Inception V3 architecture is applied and trained to identify diatoms. The architecture of employed model is shown in Fig. 37. The sensitivity and specificity are employed to evaluate the performance of proposed method. In this experiment, 43 slide images are collected for training and 10 slide images are collected for validation. Result shows that the successfully identification rate of region of interesting (ROI) reaches 89.6%.

In Qian et al. (2020), an automatic method based on Faster R-CNN is presented for algal detection. The proposed framework is depicted in Fig. 39. As seen in Fig. 38, the origin Faster R-CNN is extended based on adding extra classification branches. A data set containing 1859 samples of 37 algae in six biological categories as well as annotations of genera and classes is prepared. In this experiment, algae no more than 10 images are all classified as other genus. Therefore, 27 genera of algae is applied. 80% images of each genera is grouped into training set, while the remaining images is grouped into testing set. Result shows that the proposed method achieve 74.64% mAP on detection at genus level, and 81.17% mAP at class level.

In Baek et al. (2020), a Fast R-CNN and CNN based system is deigned to detect five cyanobacteria. The proposed system combines with classification part and counting part. The overview of proposed method in algae detection is shown in Fig. 39. Two hundred images are obtained for each of the five cyanobacteria (*Microcystis aeruginosa*, *Microcystis wesenbergii*, *Dolichospermum*, *Oscillatoria*, and *Aphanizomenon*). The precision-recall index is applied for evaluating classification performance, while the coefficient of determination and root-meansquared error are selected to determine the
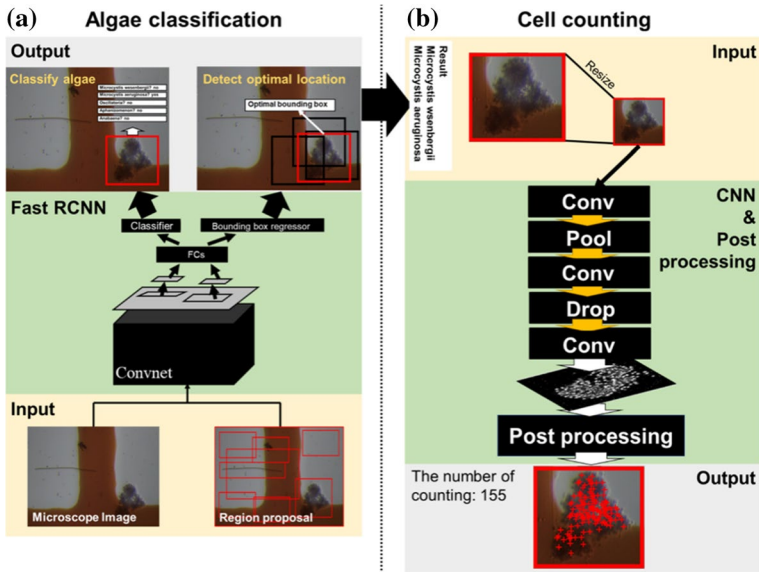
Fig. 39 The overview of proposed method mentioned in Baek et al. (2020). The figure corresponds to Fig. 2 in the original paper

accuracy of the cell counting performance. Results show that the proposed model yields the highest AP values of 0.973 on *Microcystis wesenbergii* of five cyanobacteria.

In Kang et al. (2020), a hybrid deep learning framework named as "Fusion-Net" is proposed for foodborne pathogens detection, consisting of long-short term memory network, ResNet and 1D-CNN. The Fusion-Net is generated by hyperparameter optimization, multiple deep learning architecture selection, and Fusion-Net construction. The framework of Fusion-Net is shown in Fig. 40. 5000 bacterial cells of five common foodborne bacterial cultures are prepared, which are randomly grouped into training dataset (72%), validation dataset (18%) and test dataset (10%). Based on Fusion-Net, the classification accuracy is improved up to 98.4%.

In Salido et al. (2020), for choosing the most suitable algorithm for diatom detection, the performances of YOLO and SegNet are compared. The prepared dataset contains 80 species of diatoms, each of which contains dozens to hundreds of images. Through comprehensive comparison and analysis of the specificity, sensitivity and precision of the detection results of the two models, The YOLO model is finally selected and integrated into the microscope software to provide a live diatom detection.

In Ruiz-Santaquiteria et al. (2020), the performances of semantic segmentation and instance segmentation in detecting diatom are compared. SegNet is selected as the semantic segmentation model to detect diatom. Mask R-CNN is selected as instance segmentation model. A total of 126 images of ten different taxa are analysed. Among them, 105 samples are for training and others are for validation. Compare with semantic segmentation method, the proposed instance segmentation achieves a better performance with a sensitivity of 85% and a specificity of 91%. In addition, without taking into account the reduction in specificity and precision, semantic segmentation achieves an average sensitivity of 95%.
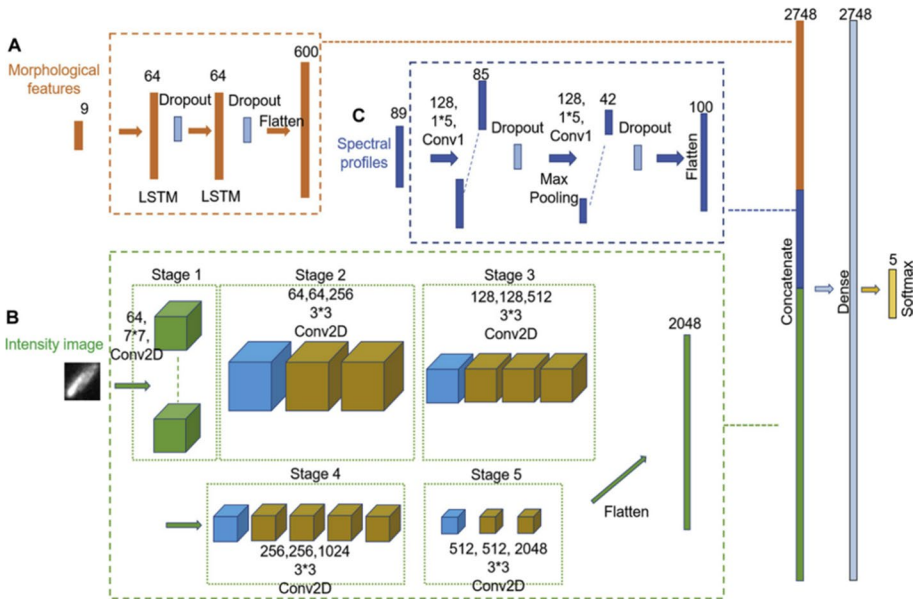
Fig. 40 The architecture of Fusion-Net mentioned in Kang et al. (2020). The figure corresponds to Fig.2 in the original paper

5.3 Summary

With the rapid development of deep learning, more and more researches employ deep learning methods for microorganism detection. The related works are shown in Table 6, where the publication date, literature links, research objects, main methods and evaluation indicators are displayed.

6 Methodology analysis

In this section, A deep analysis of classical image processing based methods, traditional machine learning based methods and deep learning based methods are respectively compiled in this section. Then, methods that are not limited to the visual transformer can be employed directly or indirectly in microorganism detection. After that, other research fields, where detection methods mentioned in this review have potential to be employed in, are introduced and analysed. Finally a small summary of this section is presented.

6.1 Analysis of common methods

In this subsection, further analysis of classical image processing based methods, traditional machine learning based methods and deep learning based methods are compiled.

Table 6 Summary for object detection method based on deep learning

Date	References	Microorganism	Method	Evaluation
2016	Akintayo et al. (2016)	Soybean cyst nematode eggs	An end-to-end convolutional selective autoencoder	Average accuracy = 94.33%
2017	Hung and Carpenter (2017)	RBC	Faster RCNN, ALexNet	Total accuracy = 98%.
2018	Tahir et al. (2018)	Fungus	CNN	Accuracy = 94.8%
2018	Panicker et al. (2018)	TB	Otsu, CNN	Recall = 97.13%, Precision =78.4%, F-score = 86.76%
2018	Pedraza et al. (2018)	Diatoms	R-CNN, YOLO	F-measure of YOLO = 84%
2019	Viet et al. (2019)	Human parasite eggs	Faster R-CNN	MAP = 97.67%
2019	Treebupachatsakul et al. (2016)	Bacteria	CNN	Accuracy = 96%
2019	Sajedi et al. (2019)	Actinobacterial species	CNN, ResNet, transfer learning	Accuracy
2019	Zhou et al. (2019)	Diatoms	Inception V3	Identification rate = 89.6%.
2020	Qian et al. (2020)	Algal	Faster R-CNN	MAP
2020	Baek et al. (2020)	Cyanobacteria	Fast R-CNN, CNN	AP
2020	Kang et al. (2020)	Foodborne pathogens	Fusion-Net	Classification accuracy = 98.4%.
2020	Salido et al. (2020)	Diatoms	YOLO, SegNet	Specificity, sensitivity, precision
2020	Ruiz-Santaquiteria et al. (2020)	Diatoms	SegNet, Mask R-CNN	Precision, sensitivity, specificity

For each category of methods, a summary chart is equipped to illustrate the commonly used methods.

6.1.1 Analysis of classical image processing based methods

According to the survey on classical image processing based methods for microorganism detection, the detection process of this methods mainly includes four steps: preprocessing, segmentation, post-processing and classification. Figure 41 shows the main processing flow and the commonly used algorithms. The most widely used segmentation algorithm for detection is threshold segmentation, which has simple calculation and high computational efficiency. The papers involved in based on threshold are Bloem et al. (1995), Baillieul and Scheunders (1998), Qing et al. (2006), Costa et al. (2008), Zhang et al. (2008), Rizvandi et al. (2008a), Rizvandi et al. (2008), Huang et al. (2008), Rizvandi et al. (2008b), Zhou and Baek (2008), Wang et al. (2008), Fernandez et al. (2008), Sotaquira et al. (2009), Zhai et al. (2010), Raof et al. (2011), Shi et al. (2012), Badsha et al. (2013), Kowalski et al. (2014), Payasi and Patidar (2017), Javidi et al. (2005), and Fernandez-Canque et al. (2006). There are many types of thresholds to choose from Sezgin and Sankur (2004). The more commonly used threshold selection method is Otsu threshold. Its main idea is selecting an optimal threshold automatically from a gray level histogram by a discriminant criterion Otsu (1979). The papers involved in based on Otsu threshold are Zhang et al. (2008), Badsha et al. (2013), Rachna and Swamy (2013), Kurtulmuş and Ulu (2014), and Goyal et al. (2015).

The calculation process of the Otsu threshold is simple and has strong generality Otsu (1979). Moreover, satisfactory results can be obtained by employing the Otsu threshold, even if the pixel values of two classes is close Sezgin and Sankur (2004). However, when the object area is much smaller than the background area, the Otsu threshold cannot provide a good segmentation result Lee and Park (1990). The large variances of the object and the background intensities and the small mean difference are responsible for the Otsu threshold to degrade its performance Lee et al. (1990).
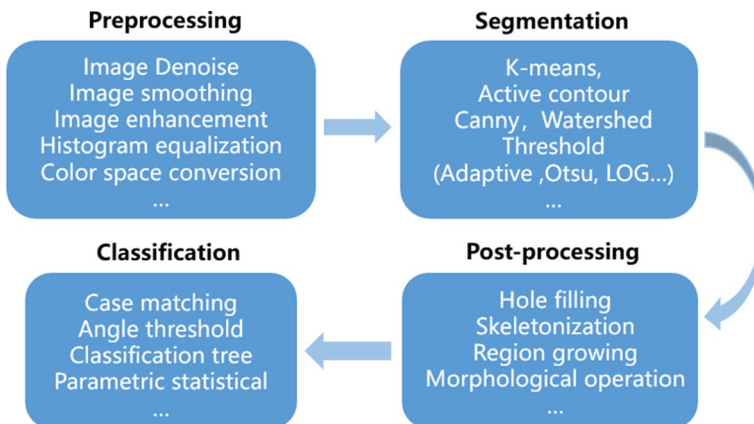

Fig. 41 Main processing flow and commonly used algorithms of the first category of detection methods

6.1.2 Analysis of traditional machine learning based methods

Common traditional machine learning method includes SVM, CRF, perceptron, *k*-NN, decision tree and logistic regression model. The main processing flow, widely used features and classification algorithms are shown in Fig. 42. According to the survey on traditional machine learning methods for microorganism detection, the most widely applied classification model is SVM, mentioned in White et al. (2010), Khutlang et al. (2010), Chang et al. (2012), Verikas et al. (2012), Li et al. (2013), Zetsche et al. (2014), Santiago-Mozos et al. (2013), Li et al. (2015), Verikas et al. (2014), and Shan e Ahmed Razaa et al. (2015). SVM is a computer algorithm that learns to assign labels to objects by example Boser et al. (1992). It constructs an optimal separating hyperplane in a higher dimensional space mapped the data into Suykens and Vandewalle (1999). In addition, the most widely extracted feature is shape feature, used in Ochoa et al. (2010), Kumar and Mittal (2010), Zhang et al. (2010), Hiremath and Bannigidad (2010), Mansoor et al. (2011), Kemmler et al. (2011), Osman et al. (2011a), Hiremath and Bannigidad (2012), Ding et al. (2012), Chang et al. (2012), Mosleh et al. (2012), Verikas et al. (2012), Zhai et al. (2012), Zetsche et al. (2014), and Verikas et al. (2014). In fact, shape feature is regarded as the most useful traditional feature in the detection of most microorganisms.

SVM can effectively use smaller training samples. This enables SVM to achieve higher classification accuracy on a smaller training set (Mercier and Lennon 2003). In addition, the idea of selecting the best hyperplane makes SVM have excellent generalization ability (Tzotsos and Argialas 2008). However, SVM is not suitable for solving multi-classification tasks (Noble 2006). Moreover, SVM is sensitive to the selection of parameters and kernel functions (Chang and Lin 2011). This means that different choices will have a great impact on the final classification accuracy of SVM.

6.1.3 Analysis of deep learning based methods

Deep learning is good at handling more complex tasks such as image recognition, object detection, speech recognition, and trend prediction. Figure 43 shows deep learning models used in microorganism detection and the publication year of each model. As one of the classical deep learning models, CNN and its derivative models are often used in the task of microorganism detection, such as CNN mentioned in Panicker et al. (2018), Tahir et al. (2018), and Sajedi et al. (2019), R-CNN mentioned in Pedraza et al. (2018), Faster R-CNN mentioned in Hung and Carpenter (2017), Viet et al. (2019), Baek et al. (2020), and Qian
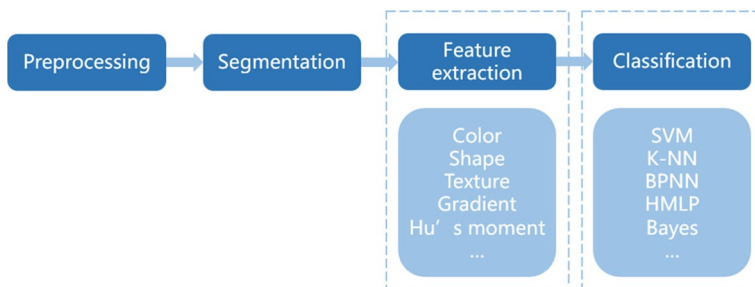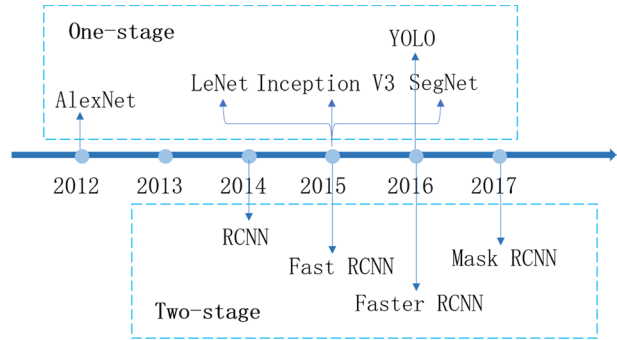

Fig. 42 Main processing flow and widely used features ang classification algorithms of the second category of detection methods

Fig. 43 Deep learning models used in microorganism detection. The horizontal arrow in the middle is the timeline. Models above time line is one-stage detection model, and the following is two-Stage detection model

et al. (2020), YOLO mentioned in Pedraza et al. (2018) and Salido et al. (2020), Mask R-CNN mentioned in Ruiz-Santaquiteria et al. (2020). Among them, Faster R-CNN is the most commonly used method. Compared to Fast R-CNN, Faster R-CNN is combined with region proposal network (RPN) to break through the bottleneck of regional proposal calculation Ren et al. (2016).

The use of RPN enables Faster R-CNN to not only improve the overall detection efficiency and accuracy, but also truly realize end-to-end detection Ren et al. (2016). Although the detection performance of Faster R-CNN is very good, it cannot perform real-time detection of object. In addition, there is a deviation in the mapping between the feature map of Faster R-CNN and the original image, which is solved by RoIAlign layer in Mask R-CNN He et al. (2017). Moreover, the RPN-based region proposal extraction method requires a large amount of calculation.

6.2 Potential methods for microorganism detection

The subsection pays attention to some new methods that have not been used in microorganism detection but have the potential. In Sect. 6.2.1, the visual transformer is firstly briefly introduced, which has received much attention from computer vision. Moreover, the feasibility of applying computer vision to microorganism detection is analysed. In Sect. 6.2.2, other potential methods are introduced and analysed.

6.2.1 Potential visual transformer methods

The transformer is firstly employed in natural language processing Devlin et al. (2018). Considering the series of achievements made by the transformer architecture in natural language processing, transformer is tried to be applied to computer vision. In fact, due to have a more robust ability of global information representation than the CNNs, it can replace the role of CNNs in vision applications. It turns out that visual transformer methods have achieved good results in many fields of computer vision, such as image segmentation (Chen et al. 2021; Liang et al. 2020), image classification (Chen et al. 2021; Liu et al. 2020), image detection (Aubreville et al. 2017; Sun et al. 2021). Although visual transformer methods show great potential, it have not yet been applied to the field of microorganism detection. Next, some Transformers that have potential applications in microorganism detection will be introduced and analyzed.
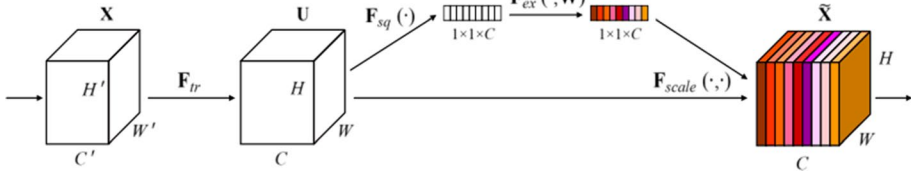
Fig. 44 The architecture of mentioned in Hu et al. (2018). The figure corresponds to Fig.1 in the original paper

In Hu et al. (2018), a new block called squeezeand-excitation (SE) based on channel domain attention mechanism is designed to recalibrate channel-wise feature responses, shown in Fig. 44. Based on SE block, a SENet is proposed with great generalization. Experiment result shows that SE is able to provide significant performance improvements for existing deep architectures with minimal additional computational cost.

In Jaderberg et al. (2015), a new learn-able module called spatial transformer based on spatial domain attention mechanism is proposed, which can perform explicit spatial transformations of features in NN. The spatial transformer module can provide an appropriate spatial transformation to each input sample. The architecture of a spatial transformer module is shown in Fig. 45. Experimental result demonstrates that the participation of spatial transformer can effectively improve the detection performance of the CNNs tested. Based on this, spatial transformer is possible to improve the performance of CNNs in microorganism detection.

In DeVries et al. (2017), a dataset augmentation method is proposed, which has nothing to do with domain. The key idea of is constructing a feature space by sequence auto-encoder and then transforming the encoded data by adding noise, interpolating or extrapolating. Experiment result shows that extrapolation is well suited for dataset augmentation in feature space. In addition, the method provides an idea of data augmentation in microorganism detection

In Beal et al. (2020), a Vision Transformer-Faster RCNN (ViT-FRCNN) is proposed to confirm the feasibility of Vision Transformer to perform object detection tasks. The key idea of ViT-FRCNN is that ViT is firstly applied to generate a a spatial feature map, which is then fed to Faster R-CNN for final detection. The architecture of ViT-FRCNN is depicted in Fig. 46. Experiment results on COCO 2017 validation set suggest that ViT-FRCNN can achieve reliable performance on the detection task. This means that ViT-FRCNN cannot only be used for microorganism detection, but also hope to achieve good results.

In Woo et al. (2018), a convolutional block attention module (CBAM) based on both channel and spatial domain attention mechanism is presented. After obtaining a feature map, CBAM respectively infers attention maps on both channel and spatial

Fig. 45 The architecture of a spatial transformer module mentioned in Jaderberg et al. (2015). The figure corresponds to Fig.2 in the original paper
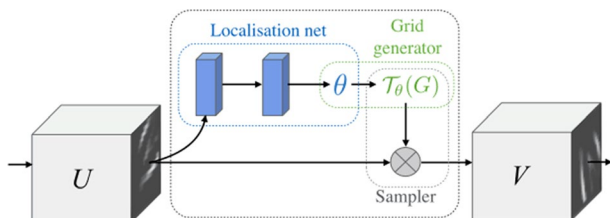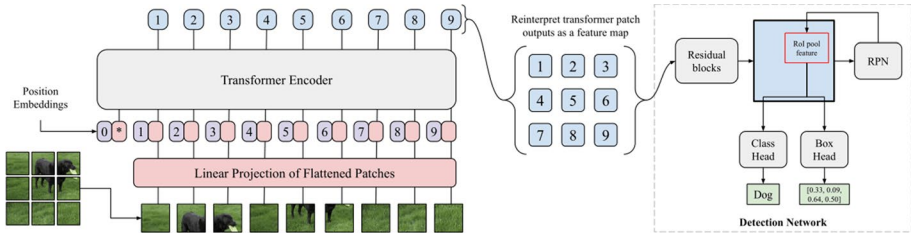
Fig. 46 The architecture of ViT-FRCNN mentioned in Beal et al. (2020).The figure corresponds to Fig.1 in the original paper

domain, which are then delivered to the input feature map. Results suggest that CBAM cannot only integrate into any CNN architectures easily, but also improve the detection performances.

In Carion et al. (2020), a novel framework called DEtection TRansformer (DETR) is proposed to achieve end-to-end object detection. The main architecture of DETR consists of a CNN backbone for feature extraction, an encoder-decoder transformer and a detection prediction network based on feed forward network. The specific framework of DETR is shown in the Fig. 47. Experiment results on COCO indicate that the results of DETR and
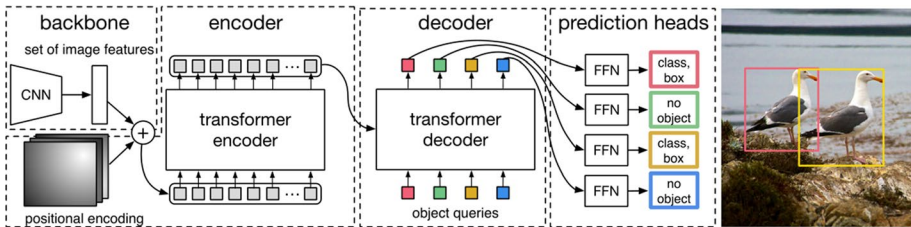

Fig. 47 The architecture of DETR mentioned in Carion et al. (2020). The figure corresponds to Fig.2 in the original paper
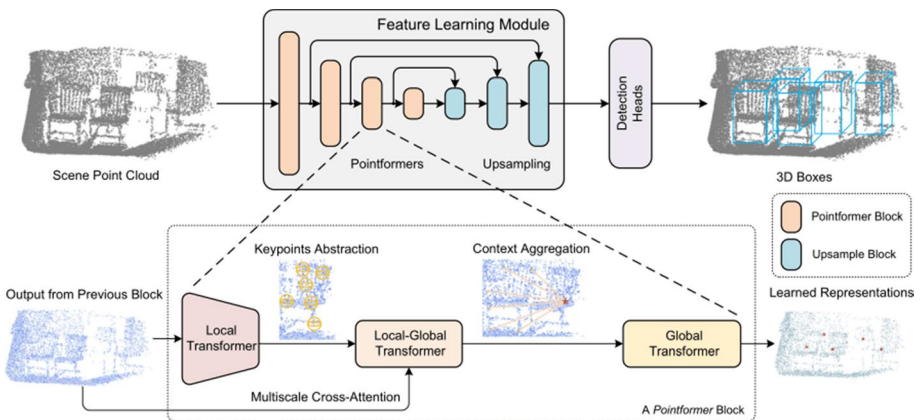

Fig. 48 The architecture of mentioned in Pan et al. (2020). The figure corresponds to Fig.2 in the original paper
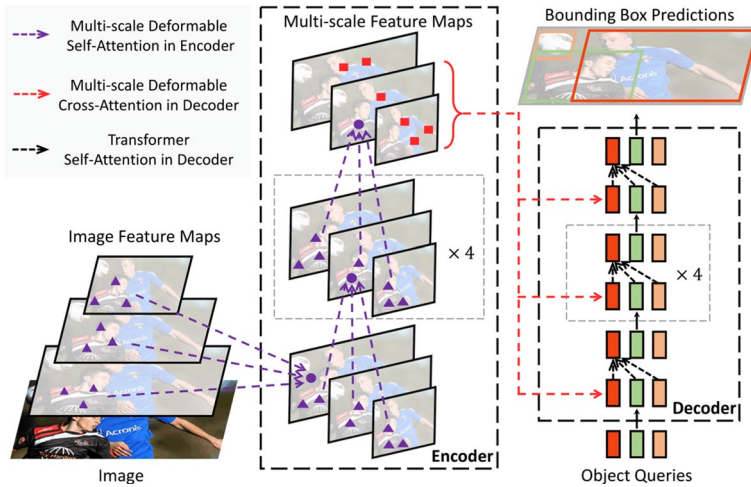
Fig. 49 The architecture of Deformable DETR mentioned in Zhu et al. (2021). The figure corresponds to Fig. 1 in the original paper
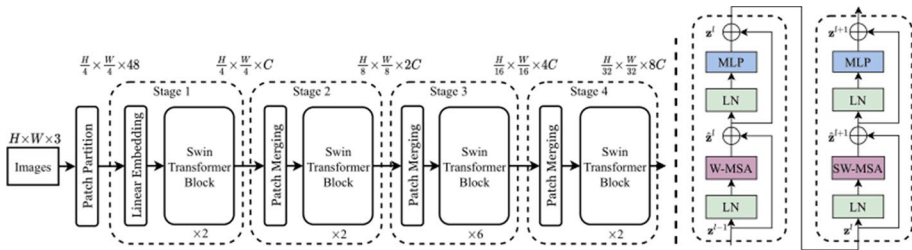

Fig. 50 Model architecture of a Swin Transformer (Swin-T) and two Swin Transformer blocks mentioned in Liu et al. (2021). The figure corresponds to Fig. 3 in the original paper

Faster R-CNN are comparable. It means DETR has the potential to be employed in microorganism detection.

In Pan et al. (2020), a Transformer backbone designed for 3-D point clouds called Pointformer is presented, which has the ability of learning features effectively. Figure 48 shows the backbone of Pointformer, which is consisted of a Local Transformer, a Local-Global Transformer and a Global Transformer. The experiment result of detection models with Pointformer show significant performance improvements. This suggests that with the assistance of Pointformer, the performance of 3-D microorganism detection is possible to improve.

In Zhu et al. (2021), to overcome the limitations of DETR's slow convergence and limited feature spatial resolution when processing image features, a Deformable DETR is presented. The framework of Deformable DETR is shown in Fig. 49. As an end-to-end object detector, the attention modules of Deformable DETR only focuses on a small set of key sampling points. This makes Deformable DETR more efficient than DETR. Experiment results on the COCO indicate that Deformable DETR is reliable in object detection. Therefore, DETR is able to be applied in microorganism detection with good performance.
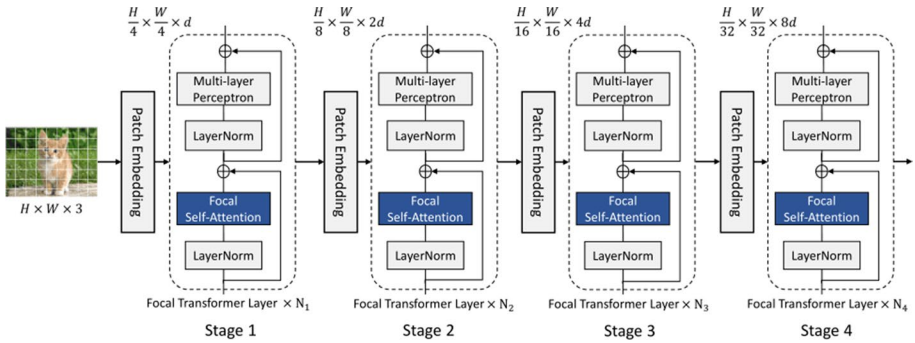
Fig. 51 The architecture of Focal Transformer mentioned in Yang et al. (2021). The figure corresponds to Fig. 2 in the original paper
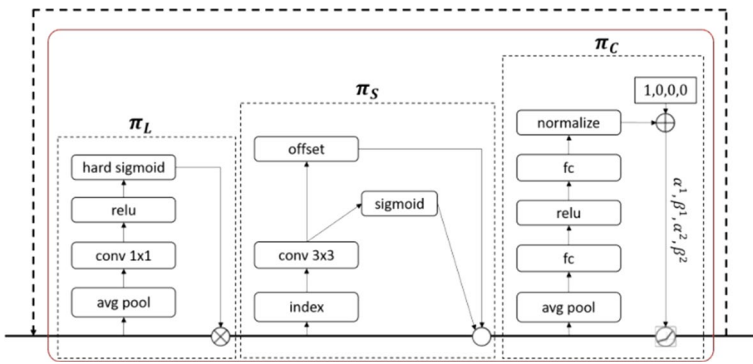

Fig. 52 The detailed implementation of dynamic head mentioned in Dai et al. (2021). The figure corresponds to Fig. 2 in the original paper

In Liu et al. (2021), a noval transformer named Swin Transformer is designed to better employ transformers applied in language field in computer vision field. Figure 50 illustrates the structure of a Swin Transformer (Swin-T) and two Swin Transformer blocks in detail. By designing a hierarchical feature calculated by Shifted windows, Swin Transformer can efficiently solve the problem of large size and high resolution of images. Experiments show that this method achieves excellent detection results on the COCO test-dev set. In Yang et al. (2021), a better transformer called Focal Transformer is proposed, which has similar structure to Swin Transformer. Unlike Swin Transformer, Focal Transformer replaces self-attention with Focal self-attention. This enables the Focal Transformer to achieve as many attention regions as the Swim Transformer at a lower cost. In addition, Focal self-attention can combine local information and global information more efficiently. Comparative experiments show that Focal Transformer can achieve better detection results than Swin Transformer (Fig. 51).

In Dai et al. (2021), for unifying object detection heads with attentions, a dynamic head framework based on combining scale-aware, spatial-aware, and task-aware attentions is proposed. The detailed implementation of dynamic head can be seen in Fig. 52.

Table 7 Summary of potential other methods for microorganism detection

Date	References	Category	Method	Potential contribution
2020	Xu et al. (2020)	Data Augmentation	EF-GANs	Extending data set
2019	Mittal et al. (2019)	Image Segmentation	B-Edge	Providing a better edge detection
2019	Sert and Avci (2019)		NS-EMFSE	Improving accuracy of segmentation
2020	Abdel-Gawad et al. (2020)		Genetic algorithm based edge detection	Providing a better edge detection
2016	Dai et al. (2016)	Detector	R-FCN	Reducing time
2016	Liu et al. (2016)		SSD	Improving precision and efficiency
2017	Li et al. (2017)		FSSD	Improving precision and efficiency
2017	Wen (2017)		II-FCN	Detecting objects in low contrast image
2017	Redmon et al. (2017)		YOLO9000	Improving accuracy
2018	Redmon and Farhadi (2018)		YOLOv3	Improving accuracy
2019	Coletta et al. (2019)		An iterative detector	Detecting new classes
2020	Liu et al. (2020)		DR-CNN	Detecting small object in a large scene
2020	Bochkovskiy et al. (2020)		YOLOv4	Improving precision and efficiency
2021	Nsaif et al. (2021)		FRCNN-GNB	Improving accuracy
2021	Qiao et al. (2021)		DetectoRS	Providing more efficient feature information
2021	Ge et al. (2021)		YOLOX	Speeding up convergence
2021	Chien-Yao et al. (2021)		YOLOR	Providing more efficient feature information

As a network module, the dynamic head can be easily combined with existing detection models to obtain better detection results.
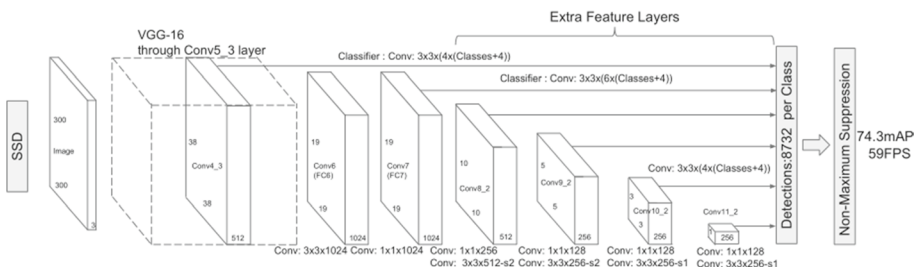

Fig. 53 The architecture of SSD mentioned in Liu et al. (2016). The figure corresponds to Fig. 2
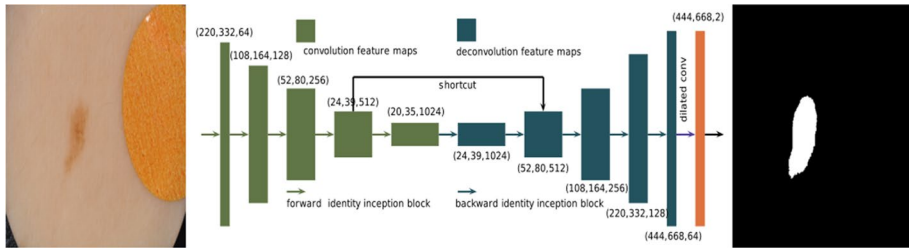
Fig. 54 The architecture of II-FCN mentioned in Wen (2017). The figure corresponds to Fig. 2 in the original paper

6.2.2 Other potential methods

For each related methods, a brief introduction is firstly made and the feasibility of used in microorganism detection is then analysed. Finally, a summary table of methods mentioned in this subsection is presented in Table 7, where the publish date, reference links, categories, main methods and potential contribution of relate works are displayed.

In Dai et al. (2016), an accurate and reliable object detection model is proposed, which is a region-based, fully convolutional networks (R-FCN). R-FCN is an improved design on the basis of ResNet-101. Compared to previous region-based detectors, R-FCN is fully convolutional, which means all required computation is shared on the entire image. The proposed method is tested on the PASCAL VOC 2007 datasets with 83.6% mAP. In addition, R-FCN requires far less detection time than Faster R-CNN. It means R-FCN has the potential to reduce the time of microorganism detection.

In Liu et al. (2016), SSD is proposed for object detection. The network architecture of SSD is shown in the Fig. 53. Compared to Faster R-CNN, the detection efficiency of SSD is higher after eliminating the process of bounding box candidate and feature up-sampling. Experimental result shows that SSD can reduce the detection time under the condition that the detection accuracy is not lower than that of Faster R-CNN. Therefore, SSD is possible to realize high precision and high efficiency detection of microorganisms.

In Shrivastava et al. (2016), an on-line hard example mining (OHEM) algorithm is proposed to train region-based ConvNet detectors, such as Faster R-CNN, Mask R-CNN. The key ides of OHEM to simplify training is automatically selecting hard examples, eliminating several common heuristics and hyper parameters. Result shows that OHEM can effectively improve the convergence of training and the final detection accuracy. With the assistance of OHEM, the performance of existing region-based ConvNet detectors in microorganism detection is able to improve.

In Li et al. (2017), a novel feature fusion single shot multibox detector (FSSD) on the basic of SSD is proooposed. Compared to SSD, FSSD enables to fuse the features easily from different scales while SSD cannot. In the feature fusion stage, FSSD makes features from different layers with different scales concatenate together. Results on the PASCAL VOC 2007 suggest that the performance of FSSD is not only superior to SSD but also superior to Faster R-CNN and YOLO. FSSD offers the possibility of obtaining better microorganism detection results.

In Wen (2017), a symmetrical identity inception fully convolution network (II-FCN) is designed to detect Lesion in Dermoscopy image, which is low contrast image. The network architecture of II-FCN is shown in Fig. 54. Results suggest that II-FCN is able to provide a accurate segmentation result even with low contrast image. It means II-FCN has
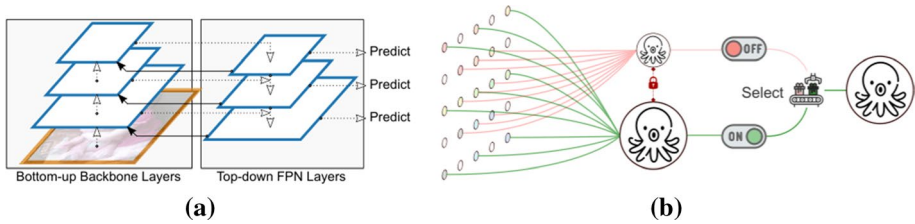
Fig. 55 a Recursive feature pyramid; **b** switchable atrous convolution. The figure corresponds to Fig. 1 in Qiao et al. (2021)

the potential to improve the performance of microorganism detection for some low contrast image.

In Coletta et al. (2019), a meaningful image detector to detect new classes that is not labelled is designed. The proposed detector is an iterative one by combining SVM and clustering algorithms. Results suggest that the presented detector has the ability of detecting new classes over time using unlabelled instances. Therefore, it provides a new idea in the detection new classes of microorganism.

In Mittal et al. (2019), a reliable edge detection algorithm based on multiple threshold approaches (B-Edge) is introduced for overcoming edge connectivity and edge thickness in edge detection. B-Edge mainly consists of three phases: graythresh and threshold computation phase, intensity adjustment phase and grayscale conversion phase. Result shows that B-Edge can detect thin edges with less noise proportion. With the assistance of B-Edge, improving the detection accuracy of microorganism is possible.

In Sert and Avci (2019), an approach called neutrosophic set-expert maximum fuzzy-sure entropy (NS-EMFSE) is proposed to for Glioblastoma, which is the most difficult in brain tumor segmentation. Experiment results prove that the performance of NS-EMFSE is batter than others. It means NS-EMFSE has the potential to improve the performance of microorganism detection by improving the accuracy of segmentation.

In Liu et al. (2020), an automatic small traffic sign detection method in large traffic scenes is proposed based on a deconvolution region-based-CNN (DR-CNN). Experimental result shows that DR-CNN can well cope with the challenge of detecting small objects in a large scene. It means when the area of the target microorganism is much smaller than the microscopic image in which it is located, the difficulty of microorganism detection can be solved by referring to the idea of DR-CNN.

In Xu et al. (2020), an enhanced framework of generative adversarial networks (EF-GANs) is proposed to deal with image augmentation problem of small datasets. EF-GANs mainly consists three steps: color space augmentation, GANs and image rotation. The experimental results suggest that the classification accuracy of corresponding classifier is improved to different degrees based on EF-GANs data expansion. This means that it is possible for EF-GANs to solve the problem of too few training samples encountered in microorganism detection, which is also a common problem in this field.

In Abdel-Gawad et al. (2020), an optimized edge detection technique is presented by applying a genetic algorithm. First, the image features are improved by balance contrast enhancement technique. After that, the fine edges is detected by combining proposed genetic algorithm and a appropriate training dataset. Through comparative experiments, the performance of the proposed method in magnetic resonance image edge detection is proved to be superior to the classical edge detection techniques. It means when applying

this method in microorganism detection, it is possible to provide a better edge detection result.

In Nsaif et al. (2021), a Faster R-CNN with Gabor filters and naive Bayes (FRCNN-GNB) model is proposed to improve the accuracy of eye detection. Compared to existing methods, the proposed model can solve the problem of occlusion or reflections from glass in eye detection. Faster R-CNN is employed to detect the initial bounding boxes of the eye region. Among this bounding boxes, Gabor filters and the naïve Bayes model are then used for determining which of them belongs to the eye region. Results suggest that the introduced algorithm is able to detect eyes with high accuracy. The idea of this method is possible to improve the accuracy of microorganism detection.

In Qiao et al. (2021), DetectoRS, a detector with improvements at both macro and micro levels, is proposed. Recursive Feature Pyramid is employed in macro level for extracting features from images that are more suitable for object detection. In addition, the employing of Switchable Atrous Convolution in micro levels enables DetectoRS to choose the training scale more flexibly and effectively. The structures of Recursive Feature Pyramid and Switchable Atrous Convolution are shown in Fig. 55. Results on COCO dataset indicate that DetectoRS yields a high box AP of 55.7% box AP for object detection.

So far, only YOLO in the YOLO series is employed in microorganism detection with good results. Compared with YOLO, YOLO9000 Redmon et al. (2017) can predict more accurately and identify more objects on the basis of ensuring that the processing speed does not decrease by applying multi-dataset joint training, new backbone network (DarkNet19) and *k*-means clustering algorithm to generate anchor box. YOLOv3 Redmon and Farhadi (2018) achieves a more accurate detection with a slight loss of detection speed by applying multi-level feature fusion and new backbone network (DarkNet53), compared with YOLO9000. YOLOv4 Bochkovskiy et al. (2020) not only achieves fast and accurate detection, but also requires less performance on the GPU during training and use. In Ge et al. (2021), YOLOX is proposed as an excellent detector. Compared with previous detectors in the YOLO series, YOLOX adopts decoupled head to speed up convergence and increase AP value of detection results. YOLOR is proposed in Chien-Yao et al. (2021) with a novel improvement perspective of combining explicit knowledge learning with implicit knowledge learning. The experimental results show that the introduction of implicit knowledge learning can improve the detection accuracy by a small margin. According to the performances of all YOLO series in addition to the original YOLO, it all have the potential to improve the performance of microorganism detection.

6.3 Potential application fields

After an in-depth study of microorganism detection methods, the reviewed methods can be applied to some other fields, such as ecosystem monitoring (Coppin et al. 2004), on-road vehicle detection (Sun et al. 2006), fabric defect detection (Ngan et al. 2011), crack detection (Mohan and Poobal 2018), forest fire detection (Alkhatib 2014), pedestrian detection (Enzweiler and Gavrila 2008), object detection in optical remote sensing images (Cheng and Han 2016), remote sensing digital image analysis (Richards and Richards 1999), histopathological image analysis (Ai et al. 2021; Chen et al. 2022; Li et al. 2022, 2021; Xue et al. 2020; Zhou et al. 2020; Li et al. 2020, 2019; Sun et al. 2020), cytopathological image analysis (Rahaman et al. 2021, 2020; Li et al. 2017), video analysis (Chen et al. 2022; Li et al. 2020; Shen et al. 2015) and COVID-19 image analysis (Rahaman et al. 2020; Li

et al. 2020). In general, the analyzed methods presented in this article provide new research ideas and better detection results for many other research fields involving detection.

7 Conclusion and future work

This review presents a state-of-the-art survey for object detection technologies in microorganism image analysis: from classical image processing and traditional machine learning to current deep learning and potential visual transformer methods. First, we introduce the basic information of microorganism detection in Sect. 1, including research motivation and research status. Second, we outline the development of early and current stages of microorganism detection methods in chronological order, including standard evaluation criteria in Sect. 2. Third, we have summarized related works in classical image processing, traditional machine learning and deep learning-based methods in Sects. 3, 4 and 5, respectively. For each category of detection methods, we present the relevant research in chronological order. For some of the outstanding research, we show its flowchart or result diagram to better understand. Finally, Sect. 6 is the method analysis part, where we analyze all the three categories of methods mentioned above and further analyze the advantages and disadvantages of some of the methods based on performance. Potential methods in microorganism detection are also introduced. Among them, methods based on visual transformer show great potential in microorganism detection.

Based on the analysis of the development of microorganism detection methods, the future development trend and challenges are predicted. The most promising development in this field can be the combination of existing detection models and visual transformer. Due to the successful applications of transformer in natural language processing, many researchers are trying to apply visual transformer to multiple computer vision tasks. Therefore, methods by combining visual transformer and others have the potential to yield good performance in the field of microorganism detection. One of the challenges can be real-time detection. With the development of related technologies, the requirements for detection speed are getting higher and higher. In many aspects such as microbial tracking, pathogen detection and wild microbial monitoring, the detection speed will significantly affect its practical application. Moreover, the difficulty of obtaining good quality microorganism data and high processing costs leads to obtaining insufficient dataset with poor quality, which encounters the experimental results. The problem of insufficient data sets and poor quality is a major challenge for microorganism detection. However, the Environmental Microorganism Image Dataset Seventh Version (EMDS-7) is available publicly Hechen et al. (2021). This dataset has 41 types of environmental microorganisms. As a multi-object dataset, it has 13216 objects in 2365 images. Furthermore, all 2365 images have corresponding labeling files in ".XML". Considering the huge development prospect in the field of microorganism detection, it is believed that more and more teams will produce high quality databases like EMDS-7.

Declarations

References

Abdel-Gawad AH, Said LA, Radwan AG (2020) Optimized edge detection technique for brain tumor detection in MR Images. IEEE Access 8:136243–136259

Adams HL, Thomas CR (1988) The use of image analysis for morphological measurements on filamentous microorganisms. Biotechnol Bioeng 32(5):707–712

Ai S, Li C, Li X, Jiang T, Grzegorzek M, Sun C, Rahaman MM, Zhang J, Yao Y, Li H (2021) A state-of-the-art review for Gastric Histopathology image Analysis approaches and future development. BioMed Res Int 2021

Akintayo A, Lee N, Chawla V, Mullaney M, Marett C, Singh A, Singh A, Tylka G, Ganapathysubramaniam B, Sarkar S (2016) An end-to-end convolutional selective autoencoder approach to Soybean Cyst Nematode eggs detection. arXiv:1603.07834

Alkhatib AAA (2014) A review on forest fire detection techniques. Int J Distrib Sensor Netw 10(3):597368

Aubreville M, Krappmann M, Bertram C, Klopfleisch R, Maier A (2017) A guided spatial transformer network for histology cell differentiation. arXiv:1707.08525

Babalola OO (2010) Beneficial bacteria of agricultural importance. Biotechnol Lett 32(11):1559–1570

Badsha S, Mokhtar N, Arof H, Lim YAL, Mubin M, Ibrahim Z (2013) Automatic cryptosporidium and giardia viability detection in treated water. EURASIP J Image Video Process 2013(1):1–10

Baek S, Pyo J, Pachepsky Y, Park Y, Ligaray M, Ahn C, Kim Y, Chun JA, Cho KH (2020) Identification and enumeration of cyanobacteria species using a deep neural network. Ecol Indic 115:106395

Baillieul M, Scheunders P (1998) On-line determination of the velocity of simultaneously moving organisms by image analysis for the detection of sublethal toxicity. Water Res 32(4):1027–1034

Beal J, Kim E, Tzeng E, Park DH, Zhai A, Kislyuk D (2020) Toward transformer-based object detection. arXiv:2012.09958

Benfield MC, Grosjean P, Culverhouse PF, Irigoien X, Sieracki ME, Lopez-Urrutia A, Dam HG, Hu Q, Davis CS, Hansen A et al (2007) RAPID: research on automated plankton identification. Oceanography 20(2):172–187

Bitton G (2005) Wastewater microbiology. Wiley, New York

Bloem J, Veninga M, Shepherd J (1995) Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. Appl Environ Microbiol 61(3):926–936

Bochkovskiy A, Wang C, Liao HM (2020) YOLOv4: optimal speed and accuracy of object detection. arXiv:2004.10934

Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on computational learning theory, pp 144–152

Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S (2020) End-to-end object detection with transformers. In: European conference on computer vision, pp 213–229. Springer, New York.

Chang C, Lin C (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2(3):1–27

Chang J, Arbeláez P, Switz N, Reber C, Tapley A, Davis JL, Cattamanchi A, Fletcher D, Malik J (2012) Automated tuberculosis diagnosis using fluorescence images from a mobile microscope. In: International conference on medical image computing and computer-assisted intervention, pp 345–352. Springer, New York

Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y (2021a) Transunet: transformers make strong encoders for medical image segmentation. arXiv:2102.04306

Chen CR, Fan Q, Panda R (2021b) Crossvit: cross-attention multi-scale vision transformer for image classification. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 357–366

Chen H, Li C, Li X, Rahaman MM, Hu W, Li Y, Liu W, Sun C, Sun H, Huang et al X (2022a) IL-MCAM: an interactive learning and multi-channel attention mechanism-based weakly supervised colorectal histopathology image classification approach. Comput Biol Med 105265

Chen A, Li C, Zou S, Rahaman MM, Yao Y, Chen H, Yang H, Zhao P, Hu W, Liu W et al (2022b) SVIA dataset: a new dataset of microscopic videos and images for computer-aided sperm analysis. Biocybern Biomed Eng

Cheng G, Han J (2016) A survey on object detection in optical remote sensing images. ISPRS J Photogramm Remote Sens 117:11–28

Chien-Yao W, I-Hau Y, Hong-Yuan ML (2021) You only learn one representation: unified network for multiple tasks

Coletta LFS, Ponti M, Hruschka ER, Acharya A, Ghosh J (2019) Combining clustering and active learning for the detection and learning of new image classes. Neurocomputing 358:150–165

Collins SP, Pope RK, Scheetz RW, Ray RI, Wagner PA, Little BJ (1993) Advantages of environmental scanning electron microscopy in studies of microorganisms. Microsc Res Tech 25(5–6):398–405

Coltelli P, Evangelisti M, Evangelista V, Gualtieri P (2007) Real-time measurement and analysis of translational and rotational speeds of moving objects in microscope fields. In: International conference on mass data analysis of images and signals in medicine, biotechnology, and chemistry, pp 128–135. Springer, Berlin

Coltelli P, Barsanti L, Evangelista V, Frassanito AM, Passarelli V, Gualtieri P (2013) Automatic and real time recognition of microalgae by means of pigment signature and shape. Environ Sci 15(7):1397–1410

Coppin P, Jonckheere I, Nackaerts K, Muys B, Lambin E (2004) Review ArticleDigital change detection methods in ecosystem monitoring: a review. Int J Remote Sens 25(9):1565–1596

Costa M, Filho C, Sena J, Salem J, Lima M (2008) Automatic identification of mycobacterium tuberculosis with conventional light microscopy. In: 2008 30th annual international conference of the IEEE engineering in medicine and biology society, pp 382–385. IEEE

Dai J, Li Y, He K, Sun J (2016) R-fcn: Object detection via region-based fully convolutional networks. arXiv:1605.06409

Dai X, Chen Y, Xiao B, Chen D, Liu M, Yuan L, Zhang L (2021) Dynamic head: unifying object detection heads with attentions. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 7373–7382

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1, pp 886–893. IEEE

Daley RJ, Hobbie JE (1975) Direct counts of aquatic bacteria by a modified epifluorescence technique 1. Limnol Oceanogr 20(5):875–882

DaneshPanah M, Zwick S, Schaal F, Warber M, Javidi B, Osten W (2010) 3D holographic imaging and trapping for non-invasive cell identification and tracking. J Display Technol 6(10):490–499

Dannemiller K, Ahmadi K, Salari E (2015) A new method for the segmentation of algae images using retinex and support vector machine. In: 2015 IEEE international conference on electro/Information technology (EIT), pp 361–364. IEEE

Dazzo FB, Niccum BC (2015) Use of CMEIAS image analysis software to accurately compute attributes of cell size, morphology, spatial aggregation and color segmentation that signify in situ ecophysiological adaptations in microbial biofilm communities. Computation 3(1):72–98

Devlin J, Chang M, Lee K, Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805

DeVries T, Taylor GW (2017) Dataset augmentation in feature space. arXiv:1702.05538

Dhindsa A, Bhatia S, Agrawal S, Sohi BS (2020) An efficient microbes detection system using microscopic images via morphological and correlation based features. Biomed Pharmacol J 13(3):1113–1124

Ding Y, Yin Y, Wang M (2012) Rapid detection based on machine vision for *Escherichia coli* in vegetables. Trans Chin Soc Agric Mach p 02

Dubuisson M, Jain AK, Jain MK (1994) Segmentation and classification of bacterial culture images. J Microbiol Methods 19(4):279–295

Embleton KV, Gibson CE, Heaney SI (2003) Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method. J Plankton Res 25(6):669–681

Enzweiler M, Gavrila DM (2008) Monocular pedestrian detection: survey and experiments. IEEE Trans Pattern Anal Mach Intell 31(12):2179–2195

Fang SP, Hsu HJ, Hung LL, Wu YS (2008) Automatic identification of Mycobacterium tuberculosis in acid-fast stain sputum smears with image processing and neural networks. Dep Electron Eng

Farahi M, Rabbani H, Talebi A, Sarrafzadeh O, Ensafi S (2015) Automatic segmentation of leishmania parasite in microscopic images using a modified cv level set method. In: Seventh international conference on graphic and image processing (ICGIP 2015), vol 9817, pp 98170K. International Society for Optics and Photonics,

Felzenszwalb P, McAllester D, Ramanan D (2008) A discriminatively trained, multiscale, deformable part model. In: 2008 IEEE conference on computer vision and pattern recognition, pp 1–8. IEEE

Fernandez H, Hintea S, Csipkes G, Pellow A, Smith H (2008) Machine vision application to the detection of micro-organism in drinking water. In: International conference on knowledge-based and intelligent information and engineering systems, pp 302–309. Springer, Berlin

Fernandez-Canque H, Beggs B, Smith E, Boutaleb T, Smith H, Hintea S (2006) Micro-organisms detection in drinking water using image processing. Cell 15:4–2

Fischler MA, Elschlager RA (1973) The representation and matching of pictorial structures. IEEE Trans Comput 100(1):67–92

Forero M, Cristobal G, Alvarez-Borreg J (2003) Automatic identification techniques of tuberculosis bacteria. In: Applications of digital image processing XXVI, vol 5203, pp 71–81. Int Soc Optic Photonics

Fukuda T, Hasegawa O (1989) Expert system driven image processing for recognition and identification of microorganisms. In: International workshop on industrial applications of machine intelligence and vision, pp 33–38. IEEE

Ge Z, Liu S, Wang F, Li Z, Sun J (2021) Yolox: exceeding yolo series in 2021. arXiv:2107.08430

Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448

Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587

Gopinath SCB, Tang T, Chen Y, Citartan M, Lakshmipriya T (2014) Bacterial detection: from microscope to smartphone. Biosens Bioelectron 60:332–342

Goyal A, Roy M, Gupta P, Dutta MK, Singh S, Garg V (2015) Automatic detection of mycobacterium tuberculosis in stained sputum and urine smear images. Arch Clin Microbiol 6(3):1

Gray TRG (1967) Stereoscan electron microscopy of soil microorganisms. Science 155(3770):1668–1670

Haoyuan C, Chen L, Xiaoyan L, Ge W, Weimin H, Yixin L, Wanli L, Changhao S, Yudong Y, Yueyang T, Marcin G (2021) GasHis-transformer: a multi-scale visual transformer approach for gastric histopathology image classification

He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37(9):1904–1916

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

He K, Gkioxari G, Dollár P, Girshick R (2017) Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969

Hechen Y, Chen L, Xin Z, Bencheng C, Jiawei Z, Pingli M, Peng Z, Ao C, Tao J, Hongzan S, Yueyang T, Shouliang Q, Tao J, Marcin G (2021) EMDS-7: environmental microorganism image dataset seventh version for multiple object detection evaluation

Hiremath PS, Bannigidad P (2010) Automatic identification and classification of bacilli bacterial cell growth phases. IJCA Spec Issue Recent Trends Image Process Pattern Recognit 1(2):48–52

Hiremath PS, Bannigidad P (2012) Spiral bacterial cell image analysis using active contour method. Int J Comput Appl 37(8):5–9

Hiremath P, Bannigidad P (2010) Digital image analysis of cocci bacterial cells using active contour method. In: 2010 international conference on signal and image processing, pp 163–168. IEEE

Hiremath PS, Bannigidad P, Hiremath M (2011) Segmentation and identification of rotavirus—a in digital microscopic images using active contour model. In: Thinkquest~ 2010, pp 177–181. Springer, Berlin

Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7132–7141

Hu W, Li C, Li X, Rahaman MM, Ma J, Zhang Y, Chen H, Liu W, Sun C, Yao Y et al (2022) GasHisSDB: a new gastric histopathology image dataset for computer aided diagnosis of gastric cancer. Comput Biol Med 105207

Huang KM, Cosman P, Schafer WR (2008) Automated detection and analysis of foraging behavior in Caenorhabditis elegans. J Neurosci Methods 171(1):153–164

Hui DS, Azhar EI, Madani TA, Ntoumi F, Kock R, Dar O, Ippolito G, Mchugh TD, Memish ZA, Drosten C et al (2019) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-the latest novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 91(264–266):2020

Hung J, Carpenter A (2017) Applying faster R-CNN for object detection on malaria images. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 56–61

Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K (2015) Spatial transformer networks. arXiv:1506.02025

Jan Z, Rafiq M, Muhammad H, Zada N (2015) Detection of tuberculosis bacteria in sputum slide image using morphological features. In: International conference: beyond databases, architectures and structures, pp 408–414. Springer, Berlin

Javidi B, Moon I, Yeom S, Carapezza E (2005) Three-dimensional imaging and recognition of microorganism using single-exposure on-line (SEOL) digital holography. Optic Express 13(12):4492–4506

Javidi B, Yeom S, Moon I, Daneshpanah M (2006) Real-time automated 3D sensing, detection, and recognition of dynamic biological micro-organic events. Optics Express 14(9):3806–3829

Javidi B, Moon I, Daneshpanah M (2010) Detection, identification and tracking of biological micro/nano organisms by computational 3D optical imaging. In: Biosensing III, vol 7759, p 77590R. International Society for Optics and Photonics

Kang R, Park B, Eady M, Ouyang Q, Chen K (2020) Single-cell classification of foodborne pathogens using hyperspectral microscope imaging coupled with deep learning frameworks. Sens Actuators B 309:127789

Kemmler M, Fröhlich B, Rodner E, Denzler J (2011) DETECTION OF MICROORGANISMS IN COMPLEX MICROSCOPY IMAGES1. TMBWK ProExzellenz project "MikroPlex" (PE113-1)

Khutlang R, Krishnan S, Whitelaw A, Douglas TS (2010) Automated detection of tuberculosis in Ziehl-Neelsen-stained sputum smears using two one-class classifiers. J Microsc 237(1):96–102

Kosov S, Shirahama K, Li C, Grzegorzek M (2018) Environmental microorganism classification using conditional random fields and deep convolutional neural networks. Pattern Recognit 77:248–261

Kovesi P (2000) Phase congruency: a low-level image invariant. Psychol Res 64(2):136–148

Kowalski M, Kaczmarek P, Kabaciński R, Matuszczak M, Tranbowicz K, Sobkowiak R (2014) A simultaneous localization and tracking method for a worm tracking system. Int J Appl Math Comput Sci 24(3):599–609

Kulwa F, Li C, Zhao X, Cai B, Xu N, Qi S, Chen S, Teng Y (2019) A state-of-the-art survey for microorganism image segmentation methods and future potential. IEEE Access 7:100243–100269

Kulwa F, Li C, Zhang J, Shirahama K, Kosov S, Zhao X, Sun H, Jiang T, Grzegorzek M (2021) A new pairwise deep learning feature for environmental microorganism image analysis. arXiv:2102.12147

Kumar S, Mittal GS (2008) Geometric and optical characteristics of five microorganisms for rapid detection using image processing. Biosyst Eng 99(1):1–8

Kumar S, Mittal GS (2010) Rapid detection of microorganisms using image processing parameters and neural network. Food Bioprocess Technol 3(5):741–751

Kurtulmuş F, Ulu TC (2014) Detection of dead entomopathogenic nematodes in microscope images using computer vision. Biosyst Eng 118:29–38

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Lee H, Park R (1990) Comments on'' An optimal multiple threshold scheme for image segmentation. IEEE Trans Syst Man Cybern 20(3):741–742

Lee SUk, Chung SY, Park RH (1990) A comparative performance study of several global thresholding techniques for segmentation. Comput Vis Graph Image Process 52(2):171–190

Lenseigne B, Brodin P, Jeon HK, Christophe T, Genovesio A (2007) Support vector machines for automatic detection of tuberculosis bacteria in confocal microscopy images. In: 2007 4th IEEE international symposium on biomedical imaging: from nano to macro, pp 85–88. IEEE

Li X, Chen C (2007) Edge detection on microbe images based on histogram equalization and soft mathematical morphology. Comput Eng Des 11:2597–2599

Li Z, Zhou F (2017) FSSD: feature fusion single shot multibox detector. arXiv:1712.00960

Li C, Shirahama K, Czajkowsk J, Grzegorzek M, Ma F, Zhou B (2013) A multi-stage approach for automatic classification of environmental microorganisms. In: Proceedings of the international conference on image processing, computer vision, and pattern recognition (IPCV), p 1. The steering committee of the world congress in computer science, computer

Li C, Shirahama K, Grzegorzek M (2015a) Application of content-based image analysis to environmental microorganism classification. Biocybern Biomed Eng 35(1):10–21

Li C, Shirahama K, Grzegorzek M (2015b) Environmental microorganism classification using sparse coding and weakly supervised learning. In: Proceedings of the 2nd international workshop on environmental multimedia retrieval, pp 9–14

Li C, Shirahama K, Grzegorzek M (2016) Environmental microbiology aided by content-based image analysis. Pattern Anal Appl 19(2):531–547

Li C, Huang X, Jiang T, Xu N (2017) Full-automatic computer aided system for stem cell clustering using content-based microscopic image analysis. Biocybern Biomed Eng 37(3):540–558

Li C, Chen H, Zhang L, Xu N, Xue D, Hu Z, Ma H, Sun H (2019a) Cervical histopathology image classification using multilayer hidden conditional random fields and weakly supervised learning. IEEE Access 7:90378–90397

Li C, Wang K, Xu N (2019b) A survey for the applications of content-based microscopic image analysis in microorganism classification domains. Artif Intell Rev 51(4):577–646

Li C, Chen H, Li X, Xu N, Hu Z, Xue D, Qi S, Ma H, Zhang L, Sun H (2020a) A review for cervical histopathology image analysis using machine vision approaches. Artif Intell Rev 53(7):4821–4862

Li X, Li C, Kulwa F, Rahaman MM, Zhao W, Wang X, Xue D, Yao Y, Cheng Y, Li J et al (2020b) Foldover features for dynamic object behaviour description in microscopic videos. IEEE Access 8:114519–114540

Li C, Kulwa F, Zhang J, Li Z, Xu H, Zhao X (2020c) A review of clustering methods in microorganism image analysis. Inf Technol Biomed 13–25

Li C, Zhang J, Kulwa F, Qi S, Qi Z (2020d) A SARS-CoV-2 microscopic image dataset with ground truth images and visual features. In: Chinese conference on pattern recognition and computer vision (PRCV), pp 244–255. Springer

Li Z, Li C, Yao Y, Zhang J, Rahaman MM, Xu H, Kulwa F, Lu B, Zhu X, Jiang T (2021a) EMDS-5: environmental microorganism image dataset fifth version for multiple image analysis tasks. PLoS ONE 16(5):e0250631

Li C, Zhang J, Rahaman MM, Yao Y, Ma P, Zhang J, Zhao X, Jiang T, Grzegorzek M (2021b) A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches

Li Y, Li C, Li X, Wang K, Rahaman MM, Sun C, Chen H, Wu X, Zhang H, Wang Q (2021c) A comprehensive review of Markov random field and conditional random field approaches in pathology image analysis. Arch Comput Methods Eng 1–31

Li X, Li C, Rahaman MM, Sun H, Li X, Wu J, Yao Y, Grzegorzek M (2022) A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches. Artif Intell Rev 1–70

Liang J, Homayounfar N, Ma W, Xiong Y, Hu R, Urtasun R (2020) Polytransform: Deep polygon transformer for instance segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 9131–9140

Lin T, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017a) Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2117–2125

Lin T, Goyal P, Girshick R, He K, Dollár P (2017b) Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp 2980–2988

Liu PY, Chin LK, Ser W, Ayi TC, Yap PH, Bourouina T, Leprince-Wang Y (2014a) An optofluidic imaging system to measure the biophysical signature of single waterborne bacteria. Lab Chip 14(21):4237–4243

Liu PY, Chin LK, Ser W, Ayi TC, Yap PH, Bourouina T, Leprince-Wang Y (2014b) Virus infectivity detection by effective refractive index using optofluidic imaging. In: 18th International conference on miniaturized systems for chemistry and life sciences, MicroTAS 2014

Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C, Berg AC (2016) Ssd: single shot multibox detector. In: European conference on computer vision, pp 21–37. Springer, Berlin

Liu Z, Li D, Ge SS, Tian F (2020a) Small traffic sign detection from large image. Appl Intell 50(1):1–13

Liu L, Hamilton W, Long G, Jiang J, Larochelle H (2020b) A universal representation transformer layer for few-shot image classification. arXiv:2006.11702

Liu W, Li C, Rahaman MM, Jiang T, Sun H, Wu X, Hu W, Chen H, Sun C, Yao Y et al (2021a) Is the aspect ratio of cells important in deep learning? A robust comparison of deep learning methods for multi-scale cytopathology cell image classification: From convolutional neural networks to visual transformers. Comput Biol Med 105026

Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B (2021b) Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 10012–10022

Locey KJ, Lennon JT (2016) Scaling laws predict global microbial diversity. Proc Natl Acad Sci 113(21):5970–5975

Mäder U, Quiskamp N, Wildenhain S, Schmidts T, Mayser P, Runkel F, Fiebich M (2015) Image-processing scheme to detect superficial fungal infections of the skin. Comput Math Methods Med

Maidak BL, Cole JR, Lilburn TG, Parker CT Jr, Saxman PR, Stredwick JM, Garrity GM, Li B, Olsen GJ, Pramanik S et al (2000) The RDP (ribosomal database project) continues. Nucl Acids Res 28(1):173–174

Mansoor H, Sorayya M, Aishah S, Mogeeb A, Mosleh A (2011) Automatic recognition system for some cyanobacteria using image processing techniques and ANN approach. Int Conf on Envir and Comp Science IPCBEE 19:73–78

Masood MI, Qadir MI, Shirazi JH, Khan IU (2011) Beneficial effects of lactic acid bacteria on human beings. Crit Rev Microbiol 37(1):91–98

Masuko M, Hosoi S, Hayakawa T (1991) A novel method for detection and counting of single bacteria in a wide field using an ultra-high-sensitivity TV camera without a microscope. FEMS Microbiol Lett 81(3):287–290

Matuszewski DJ, MLopes R (2013) Plankton detection and tracking using frequency filtering. In: 2013 Symposium on computing and automation for offshore shipbuilding, pp 75–80. IEEE

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5(4):115–133

Mercier G, Lennon M (2003) Support vector machines for hyperspectral image classification with spectral-based kernels. In: IGARSS 2003. 2003 IEEE international geoscience and remote sensing symposium. Proceedings (IEEE Cat. No. 03CH37477), vol 1, pp 288–290. IEEE

Mittal M, Verma A, Kaur I, Kaur B, Sharma M, Goyal LM, Roy S, Kim T (2019) An efficient edge detection approach to provide better edge connectivity for image analysis. IEEE Access 7:33240–33255

Mohan A, Poobal S (2018) Crack detection using image processing: a critical review and analysis. Alexandria Eng J 57(2):787–798

Moon I, Yi F, Javidi B (2010) Automated three-dimensional microbial sensing and recognition using digital holography and statistical sampling. Sensors 10(9):8437–8451

Mosleh MAA, Manssor H, Malek S, Milow P, Salleh A (2012) A preliminary study on automated freshwater algae recognition and classification system. In: BMC bioinformatics, vol 13, pp 1–13. BioMed Central

Mukti J, Kale KV, Gaikwad AT, Kulkarni A, Gadre S (2010) Detection and counting of tuberculosis bacterial cell using image processing. In: Advances and applications, computer vision and information technology, p 279

Nehl DB, Allen SJ, Brown JF (1997) Deleterious rhizosphere bacteria: an integrating perspective. Appl Soil Ecol 5(1):1–20

Ngan HYT, Pang GKH, Yung NHC (2011) Automated fabric defect detection—a review. Image Vis Comput 29(7):442–458

Noble WS (2006) What is a support vector machine? Nature Biotechnol 24(12):1565–1567

Nsaif AK, Ali SHM, Jassim KN, Nseaf AK, Sulaiman R, Al-Qaraghuli A, Wahdan O, Nayan NA (2021) FRCNN-GNB: Cascade faster R-CNN with gabor filters and naïve bayes for enhanced eye detection. IEEE Access 9:15708–15719

Nugroho HA, Akbar SA, Murhandarwati EEH (2015) Feature extraction and classification for detection malaria parasites in thin blood smear. In: 2015 2nd international conference on information technology, computer, and electrical engineering (ICITACEE), pp 197–201. IEEE

Ochoa D, Gautama S, Philips W (2010) Automatic identification of Caenorhabditis elegans in population images by shape energy features. J Microsc 238(2):173–184

Ogawa M, Tani K, Ochiai A, Yamaguchi N, Nasu M (2005) Multicolour digital image analysis system for identification of bacteria and concurrent assessment of their respiratory activity. J Appl Microbiol 98(5):1101–1106

Osman MK, Ahmad F, Saad Z, Mashor MY, Jaafar H (2010a) 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, pp 1229–1234. IEEE

Osman MK, Mashor MY, Jaafar H (2010b) Detection of mycobacterium tuberculosis in Ziehl-Neelsen stained tissue images using Zernike moments and hybrid multilayered perceptron network. In: 2010 IEEE international conference on systems, man and cybernetics, pp 4049–4055. IEEE

Osman OMK, Mashor MY, Jaafar H (2011a) Tuberculosis bacilli detection in Ziehl-Neelsen-stained tissue using affine moment invariants and Extreme Learning Machine. In: 2011 IEEE 7th international colloquium on signal processing and its applications, pp 232–236. IEEE

Osman MK, Noor MHM, Mashor MY, Jaafar H (2011b) Compact single hidden layer feedforward network for mycobacterium tuberculosis detection. In: 2011 IEEE international conference on control system, computing and engineering, pp 432–436. IEEE

Osman MK, Mashor MY, Jaafar H (2011c) Hybrid multilayered perceptron network trained by modified recursive prediction error-extreme learning machine for tuberculosis bacilli detection. In: 5th Kuala Lumpur international conference on biomedical engineering 2011, pp 667–673. Springer, Berlin

Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66

Packer HL, Thomas CR (1990) Morphological measurements on filamentous microorganisms by fully automatic image analysis. Biotechnol Bioeng 35(9):870–881

Pan X, Shi J, Luo P, Wang X, Tang X (2017) Spatial as deep: spatial cnn for traffic scene understanding. arXiv:1712.06080

Pan X, Xia Z, Song S, Li LE, Huang G (2020) 3D object detection with pointformer. arXiv:2012.11409

Panicker RO, Kalmady KS, Rajan J, Sabu MK (2018) Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng 38(3):691–699

Payasi Y, Patidar S (2017) Diagnosis and counting of tuberculosis bacilli using digital image processing. In: 2017 international conference on information, communication, instrumentation and control (ICICIC), pp 1–5. IEEE

Pedraza A, Bueno G, Deniz O, Ruiz-Santaquiteria J, Sanchez C, Blanco S, Borrego-Ramos M, Olenici A, Cristobal G (2018) Lights and pitfalls of convolutional neural networks for diatom identification. In: Optics, photonics, and digital technologies for imaging applications V, vol 10679, p 106790G. International Society for Optics and Photonics

Pepper IL, Gerba CP, Gentry TJ, Maier RM (2011) Environmental microbiology. Academic Press, Cambridge

Perner P, Perner H, Janichen S, Buhring A (2004) Recognition of airborne fungi spores in digital microscopic images. In: Proceedings of the 17th international conference on pattern recognition, 2004. ICPR 2004., vol 3, pp 566–569. IEEE

Promdaen S, Wattuya P, Sanevas N (2014) Automated microalgae image classification. Procedia Comput Sci 29:1981–1992

Puchkov E (2016) Image analysis in microbiology: a review. J Comput Commun 04:8–32

Qian P, Zhao Z, Liu H, Wang Y, Peng Y, Hu S, Zhang J, Deng Y, Zeng Z (2020) Multi-target deep learning for algal detection and classification. In: 2020 42nd annual international conference of the IEEE engineering in medicine & biology society (EMBC), pp 1954–1957. IEEE

Qiao S, Chen L, Yuille A (2021) Detectors: detecting objects with recursive feature pyramid and switchable atrous convolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10213–10224

Qing S, Wu Y, Juan J, Zhao X, Que X (2006) Application of microscopic color image processing in algae recognition and statistics. Agric Mech Res 6:199–203

Rachna HB, Swamy MSM (2013) Detection of Tuberculosis bacilli using image processing techniques. Int J Soft Comput Eng (IJSCE) 3(4)

Rahaman MM, Li C, Wu X, Yao Y, Hu Z, Jiang T, Li X, Qi S (2020a) A survey for cervical cytopathology image analysis using deep learning. IEEE Access 8:61687–61710

Rahaman MM, Li C, Yao Y, Kulwa F, Rahman MA, Wang Q, Qi S, Kong F, Zhu X, Zhao X (2020b) Identification of COVID-19 samples from chest X-Ray images using deep learning: a comparison of transfer learning approaches. J X-ray Sci Technol 28(5):821–839

Rahaman MM, Li C, Yao Y, Kulwa F, Wu X, Li X, Wang Q (2021) DeepCervix: a deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. Comput Biol Med 136:104649

Raof RAA, Mashor MY, Ahmad RB, Noor SSM (2011) Image segmentation of Ziehl-Neelsen sputum slide images for tubercle bacilli detection. Image Segm 365–378

Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271

Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. arXiv:1804.02767

Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788

Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst 28

Ren S, He K, Girshick R, Sun J (2016) Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149

Richards JA, Richards JA (1999) Remote sensing digital image analysis, vol 3. Springer, Berlin

Rizvandi NB, Pizurica A, Philips W, Ochoa D (2008a) Edge linking based method to detect and separate individual C. Elegans worms in culture. In: 2008 digital image computing: techniques and applications, pp 65–70. IEEE

Rizvandi NB, Pizurica A, Philips W (2008b) Automatic individual detection and separation of multiple overlapped nematode worms using skeleton analysis. In: International conference image analysis and recognition, pp 817–826. Springer, Berlin

Rizvandi NB, Pizurica A, Rooms F, Philips W (2008c) Skeleton analysis of population images for detection of isolated and overlapped nematode C. elegans. In: 2008 16th European signal processing conference, pp 1–5. IEEE

Rose A (1974) Vision: human and electronic. Physics Today

Ruiz-Santaquiteria J, Bueno G, Deniz O, Vallez N, Cristobal G (2020) Semantic versus instance segmentation in microscopic algae detection. Eng Appl Artif Intell 87:103271

Sajedi H, Mohammadipanah F, Rahimi SAH (2019) Actinobacterial strains recognition by machine learning methods. Multimedia Tools Appl 78(14):20285–20307

Salido J, Sánchez C, Ruiz-Santaquiteria J, Cristóbal G, Blanco S, Bueno G (2020) A low-cost automated digital microscopy platform for automatic identification of diatoms. Appl Sci 10(17):6033

Santiago-Mozos R, Pérez-Cruz F, Madden MG, Artés-Rodríguez A (2013) An automated screening system for tuberculosis. IEEE J Biomed Health Informatics 18(3):855–862

Schaal F, Warber M, Zwick S, van der KH, Haist T, Osten W (2009) Marker-free cell discrimination by holographic optical tweezers. J Eur Optic Soc-Rapid Publ 4

Schaap A, Rohrlack T, Bellouard Y (2012) Lab on a chip technologies for algae detection: a review. J Biophoton 5(8–9):661–672

Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y (2013) Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv:1312.6229

Sert E, Avci D (2019) Brain tumor segmentation using neutrosophic expert maximum fuzzy-sure entropy and other approaches. Biomed Signal Process Control 47:276–287

Sezgin M, Sankur B (2004) Survey over image thresholding techniques and quantitative performance evaluation. J Electron Imaging 13(1):146–165

Shah M, Mishra S, Sarkar M, Sudarshan S (2016) Automatic detection and classification of tuberculosis bacilli from camera-enabled Smartphone microscopic images. In: 2016 fourth international conference on parallel, distributed and grid computing (PDGC), pp 287–290. IEEE

Shan e Ahmed Razaa MQ, Marjanb MA, Farhana Buttc FS, Rajpoota NM (2015) Anisotropic tubular filtering for automatic detection of acid-fast bacilli in digitized microscopic images of Ziehl-Neelsen stained sputum smear samples. Progr Biomed Optics Imaging

Shen M, Li C, Huang W, Szyszka P, Shirahama K, Grzegorzek M, Merhof D, Deussen O (2015) Interactive tracking of insect posture. Pattern Recognit 48(11):3560–3571

Shi H, Shi Y, Yin Y (2012) Food bacteria auto identification method based on image treatment. J Jilin Univ (Eng Ed) 42(04):1049–1053

Shin D, Daneshpanah M, Anand A, Javidi B (2010) Optofluidic system for three-dimensional sensing and identification of micro-organisms with digital holographic microscopy. Optic Lett 35(23):4066–4068

Shrivastava A, Gupta A, Girshick R (2016) Training region-based object detectors with online hard example mining. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 761–769

Sieracki ME, Johnson PW, Sieburth JM (1985) Detection, enumeration, and sizing of planktonic bacteria by image-analyzed epifluorescence microscopy. Appl Environ Microbiol 49(4):799–810

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556,

Sklarczyk C, Perner H, Rieder H, Arnold W, Perner P (2007) Image acquisition and analysis of hazardous biological material in air. In: International conference on mass data analysis of images and signals in medicine, biotechnology, and chemistry, pp 1–14. Springer, Berlin

Son J, Baek M, Cho M, Han B (2017) Multi-object tracking with quadruplet convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5620–5629

Sotaquira M, Rueda L, Narvaez R (2009) Detection and quantification of bacilli and clusters present in sputum smear samples: a novel algorithm for pulmonary tuberculosis diagnosis. In: 2009 international conference on digital image processing, pp 117–121. IEEE

Steger C (1998) An unbiased detector of curvilinear structures. IEEE Trans Pattern Anal Mach Intell 20(2):113–125

Sun Z, Bebis G, Miller R (2006) On-road vehicle detection: a review. IEEE Trans Pattern Anal Mach Intell 28(5):694–711

Sun C, Li C, Zhang J, Rahaman MM, Ai S, Chen H, Kulwa F, Li Y, Li X, Jiang T (2020) Gastric histopathology image segmentation using a hierarchical conditional random field. Biocybern Biomed Eng 40(4):1535–1555

Sun Z, Cao S, Yang Y, Kitani KM (2021) Rethinking transformer-based set prediction for object detection. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 3611–3620

Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9(3):293–300

Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9

Tahir MW, Zaidi NA, Rao AA, Blank R, Vellekoop MJ, Lang W (2018) A fungus spores dataset and a convolutional neural network based approach for fungus detection. IEEE Trans Nanobiosci 17(3):281–290

Thiel S, Wiltshire RJ (1995) The automated detection of cyanobacteria using ddigital image processing techniques. Environ Int 21(2):233–236

Treebupachatsakul T, Poomrittigul S (2019) Bacteria classification using image processing and deep learning. In: 2019 34th international technical conference on circuits/systems, computers and communications (ITC-CSCC), pp 1–3. IEEE

Tripathi A, Jabbour RE, Treado PJ, Neiss JH, Nelson MP, Jensen JL, Snyder AP (2007) Detection and identification of a water mixture of E. coli cells and B. subtilis spores with Raman chemical imaging microscopy. In: Chemical and biological sensing VIII, vol 6554, p 65540J. International Society for Optics and Photonics

Tzotsos A, Argialas D (2008) Support vector machine classification for object-based image analysis. In: Object-based image analysis, pp 663–677. Springer, Berlin

Umbaugh SE (2005) Computer imaging: digital image analysis and processing. CRC Press, Boca Raton

Vallotton P, Mililli L, Turnbull L, Whitchurch C (2010) Segmentation of dense 2D bacilli populations. In: 2010 international conference on digital image computing: techniques and applications, pp 82–86. IEEE

Van DA, Salim AH, Cooreman E, Hossain MA, Rema A, Chambugonj N, Hye M, Kawria A, Declercq E et al (2002) Optimal tuberculosis case detection by direct sputum smear microscopy: how much better is more? Int J Tuberc Lung Dis 6(3):222–230

Verikas A, Gelzinis A, Bacauskiene M, Olenina I, Olenin Sj, Vaiciukynas E (2012a) Phase congruency-based detection of circular objects applied to analysis of phytoplankton images. Pattern Recognit 45(4):1659–1670

Verikas A, Gelzinis A, Bacauskiene M, Olenina I, Olenin S, Vaiciukynas E (2012b) Automated image analysis-and soft computing-based detection of the invasive dinoflagellate Prorocentrum minimum (Pavillard) Schiller. Expert Syst Appl 39(5):6069–6077

Verikas A, Gelzinis A, Bacauskiene M, Olenina I, Vaiciukynas E (2014) An integrated approach to analysis of phytoplankton images. IEEE J Oceanic Eng 40(2):315–326

Viet NQ, ThanhTuyen DT, Hoang TH (2019) Parasite worm egg automatic detection in microscopy stool image based on Faster R-CNN. In: Proceedings of the 3rd international conference on machine learning and soft computing, pp 197–202

Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, vol 1, pp I–I. IEEE

Wang J, Trubuil A, Graffigne C (2001) 3D biological object detection and labeling in multidimensional microscopy imaging. In: Proceedings 11th international conference on image analysis and processing, pp 215–220. IEEE

Wang J, Graffigne C, Trubuil A, Kaeffer B (2003) 3-D aggregated object detection and labeling from multivariate confocal microscopy images: a model validation approach. IEEE Trans Syst Man Cybern B 33(4):572–581

Wang P, Wen C, Li W, Chen Y (2008) Motile microorganism tracking system using micro-visual servo control. In: 2008 3rd IEEEiInternational conference on nano/micro engineered and molecular systems, pp 178–182. IEEE

Wang Y, Tao S, Chen J (2013) Image fusion technology application in water quality monitoring based on digital microscopic. In: Applied mechanics and materials, vol 295, pp 909–914. Trans Tech Publ

Wen H (2017) II-FCN for skin lesion analysis towards melanoma detection. arXiv:1702.08699

White AG, Cipriani PG, Kao H, Lees B, Geiger D, Sontag E, Gunsalus KC, Piano F (2010) Rapid and accurate developmental stage recognition of C. elegans from high-throughput image data. In: 2010 IEEE computer society conference on computer vision and pattern recognition, pp 3089–3096. IEEE

Widmer KW, Srikumar D, Pillai SD (2005) Use of artificial neural networks to accurately identify Cryptosporidium oocyst and Giardia cyst images. Appl Environ Microbiol 71(1):80–84

Woo S, Park J, Lee J, Kweon IS (2018) Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV), pp 3–19

Xu H, Li C, Rahaman MM, Yao Y, Li Z, Zhang J, Kulwa F, Zhao X, Qi S, Teng Y (2020) An enhanced framework of generative adversarial networks (EF-GANs) for environmental microorganism image augmentation with limited rotation-invariant training data. IEEE Access 8:187455–187469

Xu W, Ying J (2014) Detection of tuberculosis in sputum smear images by Gaussian mixture models. Comput Knowl Technol p 10

Xue D, Zhou X, Li C, Yao Y, Rahaman MM, Zhang J, Chen H, Zhang J, Qi S, Sun H (2020) An application of transfer learning and ensemble learning techniques for cervical histopathology image classification. IEEE Access 8:104603–104618

Yang C, Li C, Tiebe O, Shirahama K, Grzegorzek M (2014) Shape-based classification of environmental microorganisms. In: 2014 22nd international conference on pattern recognition, pp 3374–3379. IEEE

Yang J, Li C, Zhang P, Dai X, Xiao B, Yuan L, Gao J (2021) Focal self-attention for local-global interactions in vision transformers. arXiv:2107.00641

Yeom S, Moon I, Javidi B (2006) Real-time 3-D sensing, visualization and recognition of dynamic biological microorganisms. Proc IEEE 94(3):550–566

Yin Y, Ding Y (2009) Rapid method for enumeration of total viable bacteria in vegetables based on computer vision. Trans CSAE 25(7):249–254

Yourassowsky C, Dubois F (2014) High throughput holographic imaging-in-flow for the analysis of a wide plankton size range. Opt Express 22(6):6661–6673

Yu JQ, Huang W, Chin LK, Lei L, Lin ZP, Ser W, Chen H, Ayi TC, Yap PH, Chen CH et al (2014) Droplet optofluidic imaging for λ-bacteriophage detection via co-culture with host cell *Escherichia coli*. Lab Chip 14(18):3519–3524

Zetsche E, Mallahi AE, Dubois F, Yourassowsky C, Kromkamp JC, Meysman FJR (2014) Imaging-in-flow: digital holographic microscopy as a novel tool to detect and classify nanoplanktonic organisms. Limnol Oceanogr 12(11):757–775

Zhai Y, Liu Y, Zhou D, Liu S (2010) Automatic identification of mycobacterium tuberculosis from ZN-stained sputum smear: algorithm and system design. In: 2010 IEEE international conference on robotics and biomimetics, pp 41–46. IEEE

Zhai Y, Zhou D, Liu Y (2012) Recognition of mycobacterium tuberculosis in microscopic images based on color and gradient feature. J Natl Univ Defense Technol 34(5):146–152

Zhang C, Chen W, Liu W, Chen C (2008) An automated bacterial colony counting system. In: 2008 IEEE international conference on sensor networks, ubiquitous, and trustworthy computing (sutc 2008), pp 233–240. IEEE

Zhang R, Zhao S, Jin Z, Yang N, Kang H (2010) Application of SVM in the food bacteria image recognition and count. In: 2010 3rd international congress on image and signal processing, vol 4, pp 1819–1823. IEEE

Zhang J, Li C, Kulwa F, Zhao X, Sun C, Li Z, Jiang T, Li H, Qi S (2020) A multiscale CNN-CRF framework for environmental microorganism image segmentation. BioMed Res Int

Zhang J, Li C, Kosov S, Grzegorzek M, Shirahama K, Jiang T, Sun C, Li Z, Li H (2021a) LCU-Net: a novel low-cost U-Net for environmental microorganism image segmentation. Pattern Recognit 115:107885

Zhang J, Li C, Rahaman MM, Yao Y, Ma P, Zhang J, Zhao X, Jiang T, Grzegorzek M (2021b) A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches. Artif Intell Rev 1–70,

Zhao P, Li C, Rahaman MM, Yang H, Jiang T, Grzegorzek M (2021) A comparison of deep learning classification methods on small-scale image data set: from convolutional neural networks to visual transformers. arXiv:2107.07699

Zhou BT, Baek JH (2008) Using machine vision to detect distinctive behavioral phenotypes of thread-shape microscopic organism. In: Applications of computational intelligence in biology, pp 161–182. Springer, Berlin

Zhou F, Liu J (2016) Microbial contour extraction based on edge detection. In: 2016 8th international conference on wireless communications & signal processing (WCSP), pp 1–5. IEEE

Zhou Y, Zhang J, Huang J, Deng K, Zhang J, Qin Z, Wang Z, Zhang X, Tuo Y, Chen L et al (2019) Digital whole-slide image analysis for automated diatom test in forensic cases of drowning using a convolutional neural network algorithm. Forensic Sci Int 302:109922

Zhou X, Li C, Rahaman MM, Yao Y, Ai S, Sun C, Wang Q, Zhang Y, Li M, Li X et al (2020a) A comprehensive review for breast histopathology image analysis using classical and deep neural networks. IEEE Access 8:90931–90956

Zhou Y, Cao Y, Huang J, Deng K, Ma K, Zhang T, Chen L, Zhang J, Huang P (2020b) Research advances in forensic diatom testing. Forensic Sci Res 1–8

Zhu X, Su W, Lu L, Li B, Wang X, Dai J (2021) Deformable detr: deformable transformers for end-to-end object detection

Zou Y, Li C, Shirahama K, Jiang T, Grzegorzek M (2016) Environmental microorganism image retrieval using multiple colour channels fusion and particle swarm optimisation. In: 2016 IEEE international conference on image processing (ICIP), pp 2475–2479. IEEE

Zou Y, Li C, Shirahama K, Jiang T, Grzegorzek M (2017) Content-based image retrieval of environmental microorganisms using double-stage optimisation-based fusion. Inf Eng Express 3(4):43–53

Zou Z, Shi Z, Guo Y, Ye J (2019) Object detection in 20 years: a survey. arXiv:1905.05055