International Journal of Artificial Intelligence,
ISSN 0974-0635; Int. J. Artif. Intell.
2025, Volume 23, Issue Number 1

On the Task of Classifying Microorganisms for Microscopic Analysis

Mikhail Gorodnichev, Maxim Gromov, Marina Moseva and Kamil Kharrasov

Department of Mathematical Cybernetics and Information Technology, MTUCI,
8a, Aviamotornaya st., Moscow, Russia
m.g.gorodnichev@mtuci.ru, it@mtuci.ru, m.s.moseva@mtuci.ru, k.r.harrasov@edu.mtuci.ru

ABSTRACT

The study of microorganisms with a microscope is a task in demand all over the globe, in thousands of bio-laboratories for very different purposes. It can be a simple scientific study of microbial behavior, as well as the study of water quality for the content of pathogenic bacteria, highly demanded in the sanitary examination of various water basins. These and many other tasks require one common action - classification of the microorganism (correct identification of the species) in order to continue the work performed. This step requires highly qualified personnel, whose assistance may not always be possible due to limitations in both personnel resources and time. Therefore, this paper will consider the possibility of replacing the human (full or partial) to simplify the classification process. In this paper, an algorithm for microorganism classification based on YOLOv5 and Mask-RCNN is proposed. A study has been carried out on existing datasets as well as on the dataset collected by the authors. From the data available in public domain, by applying standard algorithms on data augmentation on microorganism search scenarios on image the algorithms used give good result on both detection and segmentation (later applicable for biomass counting). But this result is possible only if there are similar data format and the required microorganisms in the training sample.

Keywords: object classification, object detection, microscopic view, microorganism classification, microscopic image, transfer learning.

Mathematics Subject Classification: 68T45 Machine vision and scene understanding.

1. INTRODUCTION

Microbial research is a demanded task in thousands of bio-laboratories around the world, for many different purposes. It may be a simple scientific study of microbial behavior, or it may be a water quality study of pathogenic bacteria, which is highly demanded in sanitary examinations of various water basins. These and many other tasks require fulfillment of one important function - classification of microorganisms (correct identification of species), as identification of a specific species allows to conduct better research work. Identifying a specific microbial species is a task that requires the involvement of highly qualified personnel, which is not always possible due to both human and time constraints.

Despite the fact that AI has started to be actively applied in medicine, biology and other natural sciences, there are few examples of integrating computer vision to solve the tasks of studying and identifying microorganisms. Most often, these are studies or works that were aimed at solving a specific problem, for example, finding a certain type of microorganism on images taken on a microscope of a certain manufacturer. The results of such work are rather limited and cannot be used in further studies if another type of microscope is used. This happens because of a number of problems:

1. The problem of collecting a large amount of data needed to train a neural network. In order to get proper results of modern models for one class of study, thousands of different photos taken with different equipment are needed. Finding and capturing a particular species of microorganism is a time consuming task as researchers will have to spend time collecting material, creating incubatory conditions for the samples, and then conducting a long and diligent analysis to find the right species of microorganism manually, using a microscope.

2. Different equipment with different techniques. As it was said above, to create a really good and universal model, you need a high-quality capture of microorganisms, which, depending on the equipment of the laboratory is performed on different equipment. Various microscopes, several types of lenses, different manufacturers of lenses, brightness of illumination, camera on which the shooting is carried out, quality of reagents used, all these factors affect the quality of the image created. At the same time, unifying the requirements for equipment and image quality in order to create a diverse, large and high-quality data set is an almost impossible task due to the different equipment used in laboratories.

3. It is necessary to have highly qualified specialists to perform all the required work. Note that an essential part before training a neural network is the creation of the data set, namely the markup of the captured data. Since it is necessary to mark and correctly identify the species of microorganism on a huge number of photos, the specialist who must qualitatively perform the task may simply not have enough working time. Often, such tasks are performed by junior staff with selective checking of the results obtained by senior qualified staff, or such work is outsourced, which often leads to errors in the process of data markup.

All the problems described above lead to the emergence of completely different in structure and content datasets and models, which are obtained at the output are highly specialized and, most likely, are not applicable in practice. For example, we can cite the work of Chinese EMDS-7 specialists. Scientists have collected 42 species of microorganisms by the seventh version of the dataset, but with the quantity, the quality of the dataset has suffered greatly. There is a large number of errors in the markup - microorganisms are signed incorrectly in some photos, unrepresentative sampling - the class of one microorganism is much larger in number than all the others, because of which there may be problems in training. The photos themselves are made with a special green filter, making it impossible to use these photos to train models that will be used on microscopes without this filter. All of this is just a small part of the problems presented in this paper alone. A process such as standardization of datasets for training neural networks in microbiology would help to solve this problem.

This paper is divided into 5 sections. Section 2 provides a description of the current state of the problem. Studies on similar topics with the main results of the authors are given. Section 3 describes the main materials and methods used in the work. The main microbial datasets are reviewed, and the datasets selected for the study. A description of the neural network architectures used in the work is given. The data preprocessing process is described. Section 4 contains a description of the experiment conducted and the main results of the work. Section 5 contains the final conclusions of the work.

2. RELATED WORK

Currently there are attempts to improve the performance of microorganism recognition algorithms, so, recently many researchers have begun to apply the method of detection of specific microorganisms and bacteria using an improved algorithm YOLOv5, which is based on the introduction of a mechanism for the detection of specific specified objects, so this algorithm allows the implementation of an approach to target detection of cells and microorganisms. The approach is based on the technical implementation of Efficient Channel Attention (ECA) module which is added to the YOLOv5 model for key feature extraction, also the authors of the study replaced the path aggregation network (PANet) of YOLOv5 with bi-directional pyramidal feature network (BiFPN) for fast multi-scale feature aggregation to find specific microorganisms. The results of the study showed a detection accuracy of 81.98% for a specific microorganism, which is quite high, however, when a quick reaction is needed when a pathogen is detected in water, it may be necessary to double-check the results, which may waste precious time for decision making to minimize the damage [1].

Also, one of the microorganism detection methods is a hybrid method based on Inception-V3 and SVM mechanisms. Mentioned method is based on image preprocessing with image size conversion, flipping and gray scale to RGB, followed by feature extraction using Inception-V3 DCNN model. Meanwhile, further image classification was performed using SVM support vector method. The authors used 200 images of seven bacteria for testing. The proposed approach achieved an accuracy of 96% [2]. It should be noted that this approach actually showed a high recognition result, but the disadvantage of this system is the manual image processing, which greatly slows down the process of microorganism detection.

Note that in 2020, an approach was considered and implemented which consisted of applying a long-term memory network (LSTM), a deep residual network (ResNet) and a one-dimensional CNN (1D-CNN) network to classify food microorganisms using hyperspectral microscopic imaging (HMI) technology. The researchers used the dataset processed before them, the training results showed the following accuracy: LSTM showed 92.2% accuracy, ResNet showed 93.8% accuracy, and 1D-CNN showed 96.2% accuracy, respectively [3]. All three approaches showed a high rate of accuracy in microorganisms recognition, however, the study used a preprocessed dataset, making it difficult to estimate the time and technical costs associated with the implementation of such a system.

Also, some scientists propose an approach aimed at using deep neural networks using Extreme Gradient Boosting (XGBoost), namely, one study proposed an approach that implemented the Gabor transform to identify texture features and XGBoost to classify bacteria, the recognition accuracy of the proposed approach was 90.28% [4]. The disadvantage of the proposed system is that the use of XGBoost requires the involvement of large computing power and almost manual adjustment of hyperparameters.

Another popular method applied in the detection of microorganisms and bacteria is a hybrid method based on the use of deep convolutional neural network (DCNN), which is used to create image descriptors to then use a pool encoder to create feature vectors, after that the classification problem was solved using SVM or RF. The proposed approach achieves an image recognition accuracy of 97.24% [5]. The method described above also requires high quality images, which in turn is time-consuming and requires highly skilled professionals.

In a study related to water pollutant classification, a deep learning-based model utilizing OpenCV and TensorFlow libraries based on YOLOv3 and YOLOv4 models was used. In this study, computer vision was used to set the bounding box for each type of pollutant and the model searched for a specific pollutant based on the images provided in the training dataset. The model performed well in pollutant detection but required high quality images [6].

Also, computer vision was used to detect the change in behaviour of fish in the water body to determine the water condition. To comprehensively analyze the effect of water quality on fish behaviour, the study applied multivariate feature parameters to quantify indicators including such as movement speed, turning angle, spatial standard deviation, and body colour that characterize changes in fish behaviour. A classification model based on LSTM (long short-term memory) neural network was used for fish detection and behaviour detection, Red Danio fish was used as an indicator organism, and copper sulfate solution was used as a toxic contaminant to simulate water pollution. The experimental results showed that the accuracy of water quality classification using the proposed system reached 91-93%. The disadvantage of the proposed approach is the need to understand the exact list of pollutants entering the water, with a huge number of chemical compounds capable of harming water and water bodies [7].

In another study, detection of Escherichia coli (E. coli) and Vibrio cholera (V. cholera) bacteria in the water of water bodies using CNNs was considered. About 9000 red-green-blue (RGB) microscopic images of wastewater containing stained bacteria were used as input datasets. The results showed that the bacteria were classified and counted with an accuracy ranging from 93.01% to 97%, respectively. Although the CNN performed quite well in counting bacteria for both RGB colour and grayscale models, its classification performance was satisfactory only for RGB images. Sensitivity analysis of the CNN showed that the enhancement of Gaussian noise resulted in an increase in the standard deviation, which proportionally decreased the accuracy of the CNN [8].

3. MATERIALS AND METHODS

This section will describe the materials and methods used in this paper. Initially, the datasets are described, followed by the methods used and the experiment conducted.

The study used neural networks for object detection, namely YOLOv5 series models, the input to the object detector was an image represented as a matrix that contains information about each pixel. After that, the data was sent to train the model.

3.1. Datasets

The two most popular microorganism datasets used in this work are EMDS-7 - Environmental Microorganism Detection Dataset Seventh Version [9] used for classification task and SinfNet dataset [10] used for segmentation task.

3.1.1. EMDS-7

The EMDS-7 dataset comprises 2365 images across 42 categories of microscopic environmental objects, with annotations for 13216 labeled items. The samples were collected from various lakes and rivers in Shenyang, Northeast China, by two environmental biologists from Northeastern University, China, using a 400x optical microscope between 2018 and 2019. Subsequently, in 2020–2021, four bioinformatics experts from the same university manually prepared .XML annotation files corresponding to all 2365 original images.

Unknown microorganisms and impurities are marked as “unknown”, and a total of 13216 labeled objects are obtained.

The dataset provides an overview of 41 types of microorganisms and unidentified objects. The labeled EMDS-7 image files were manually annotated. All recognizable EMs that are present in 60% or more of their own total occurrences across all images are categorized with labels corresponding to the 41 defined categories. Unknown EMs, which constitute less than 40% of their own occurrences across all images, as well as EMs outside the 41 predefined categories, are classified as unknown. Additionally, any noticeable impurities in the image backgrounds are also marked as unknown.

After careful analysis of the labeled data, the following problems were found:
- Unequal distribution of classes. The cyanobacteria class is strongly expressed in the dataset, causing a bias towards the predominant object class during training
- Errors in data markup. The problem is obvious - incorrect markup gives errors in the subsequent work of the neural network. There were errors both in inaccurate marking of object boundaries and in incorrect labeling.
- Photos are taken at different magnifications, 100x and 400x are encountered
- Images are taken with a strong green filter - such pictures are not suitable for microscopes, where there is no green background.

- Unsuitable for real training format of stored data and detailed description of microorganisms only in Chinese language

All of the above problems had to be solved in order to test the suitability of the found data in neural networks and a huge amount of time was spent on this.

3.1.2. SinfNet

Data set with the most common microorganisms, as well as validation of the created data in object detection and image segmentation scenarios, counting microorganism biomass from visual data. This work was supported by Microsoft Corporation. An example of the data presented in this set can be seen in Figure 1.

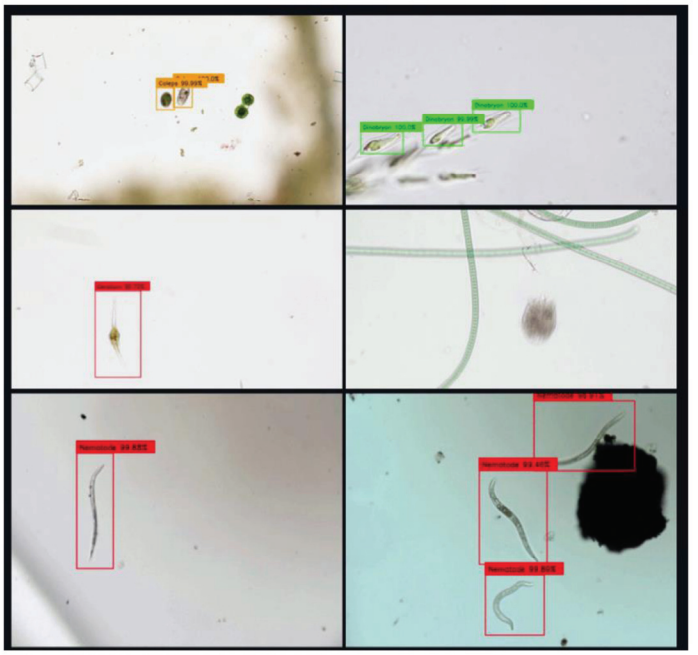

Figure 1. SinfNet data examples

Unlike the previous work, this one was done with more serious requirements to data quality, but because of this the size of the collected dataset suffered greatly. Instead of 42 classes, only 17 classes are represented.

Characteristics of the work:

- No errors were found in the markup
- In addition to protozoa, nematodes are represented
- In addition to the markup for detection, the images are qualitatively marked for segmentation
- Images on a neutral background, without filters and in very high resolution
- Tested on classical neural network models - RCNN and U-net
- Examples of biomass calculations based on detection results are presented.

3.1.3. DeepBacs

This paper [11] considers the application of neural networks in conventional light microscopy and fluorescence microscopy at ultra-high magnifications (1000x and more). The work also presents pre and post image processing scenarios for interfacing with AI algorithms. An example of the data is presented in Figure 2. The advantages of this work include:

- Error free markup
- Good result in U-net segmentation scenario
- A detailed description of the application of super-sampling (algorithms to programmatically increase image resolution) in microscopy scenarios
- But this work has some disadvantages:
- The photographs are taken in the infrared range, or fluorescence microscopy was used - rare equipment
- The images are marked up in an extremely rare format, resulting in the need to work in outdated programs, or manually repartition the data
- Tested on outdated YoloV2 detector - leading to poor detection results.

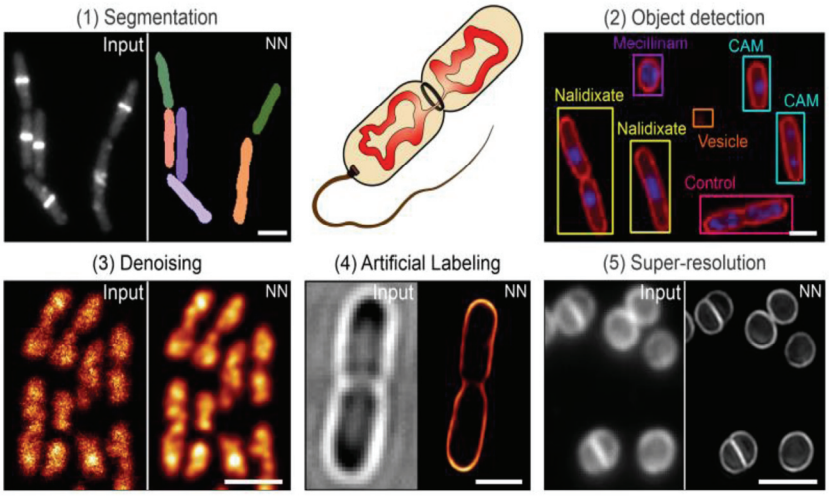

Figure 2. DeepBacs data examples

3.1.4. Other datasets

Two other useful information centers were also discovered during the study. They were not applied within the scope of this study, but in the future these databases may also be useful:

World Data Center for Microorganisms website [12] – it contains catalogs of open databases with a large list of microorganisms, their characteristics, etymology, habitats.

Digital Image of Bacterial Species (DIBaS) dataset [13] – a set of photographs of bacterial colonies. Close-up photographs of bacterial colonies. No labeling of individual microbes, different principle of bacterial classification.

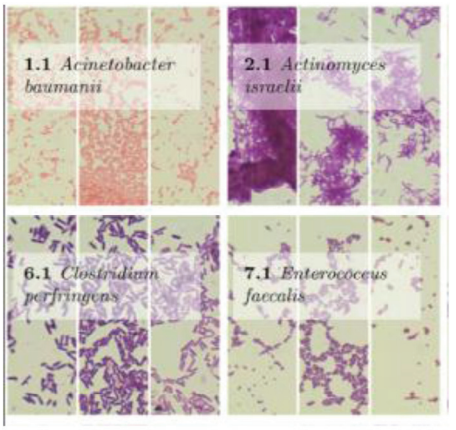

Figure 3. An example of a colony photo from the Digital Image of Bacterial Species (DIBaS) dataset

3.2. Methods
3.2.1. YOLOv5

YOLOv5 (versions 6.0 and 6.1) is an advanced algorithm for object detection. The architecture of the model is shown in the Figure 4.

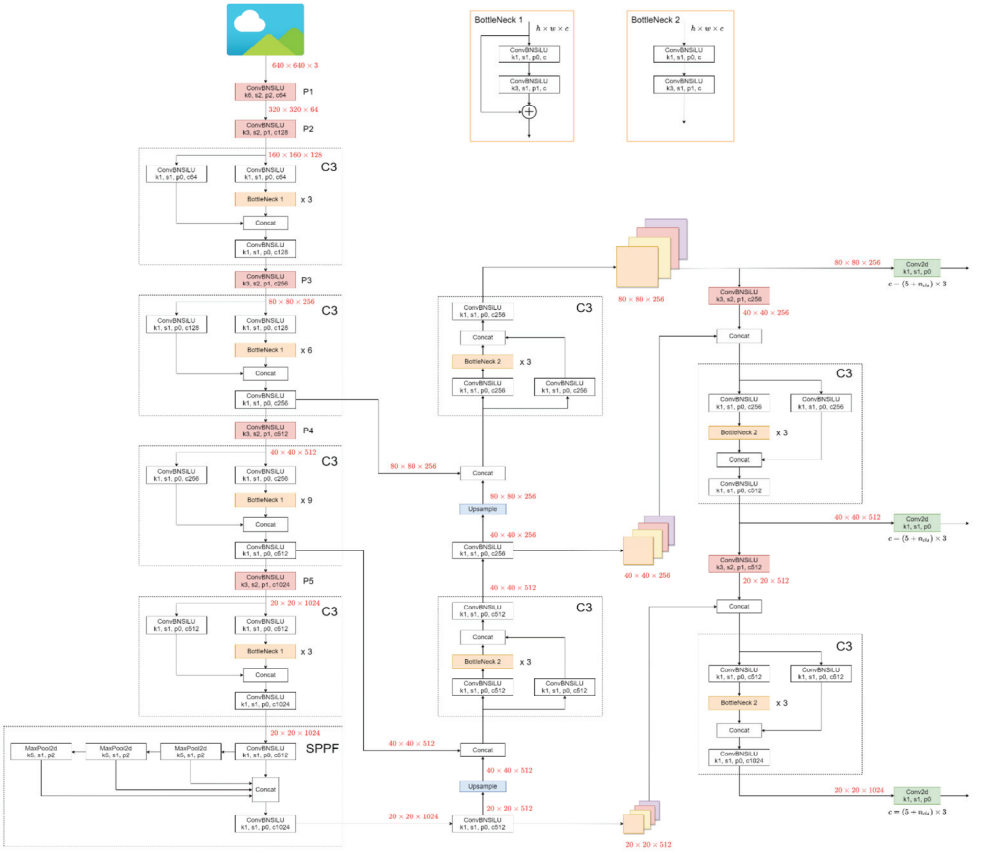

Figure 4. YOLOv5 architecture

YOLOv5 algorithm computes its loss by combining three distinct components:
- Classification Loss (BCE Loss): This binary cross-entropy loss evaluates the error associated with the classification task.
- Object Loss (BCE Loss): This binary cross-entropy loss measures the error in detecting the presence of an object within a specific grid cell.
- Localization Loss (CIoU Loss): This complete IoU loss quantifies the error in accurately positioning an object inside a grid cell
The overall loss function is depicted as:

$$Loss = \lambda_1 L_{cls} + \lambda_2 L_{obj} + \lambda_3 L_{loc} \quad (1)$$

The objectivity losses of the three prediction layers (P3, P4, P5) are weighted differently. The balance weights are [4.0, 1.0, 0.4], respectively. This approach ensures that the predictions at different scales contribute appropriately to the total loss.

$$L_{obj} = 4.0 * L_{obj}^{small} + 1.0 * L_{obj}^{medium} + 0.4 * L_{obj}^{large} \quad (2)$$

In YOLOv5, the formula for predicting box coordinates has been updated to reduce grid sensitivity and prevent the model from predicting unbounded box dimensions.

The revised formulas for calculating the predicted boundary are as follows:

$$b_x = (2 * \sigma(t_x) - 0.5) + c_x b_y = \left(2 * \sigma\left(t_y\right) - 0.5\right) + c_y b_w = p_w * \left(2 * \sigma(t_w)\right)^2 b_h = p_h * \left(2 * \sigma(t_h)\right)^2 \quad (3)$$

YOLOv5 include the use of a dynamic architecture, a wide range of data augmentation techniques, innovative learning strategies, and important adjustments to the computational loss and target construction process. All of these features lead to improve the accuracy and efficiency of object detection while maintaining the high speed that is the hallmark of YOLO models.

3.2.2. Detectron2

The Mask R-CNN architecture implemented in the Detectron2 framework was used in this paper. Detectron2 is Facebook AI Research's next generation software system that implements state-of-the-art object detection algorithms. It is a ground-up rewrite of the previous version, Detectron, and it originates from Mask R-CNN-benchmark. It consists of:

- Training recipes for object detection, instance segmentation, panoptic segmentation, semantic segmentation and key-point detection.
- 80+ pre-trained models to use for fine-tuning (or training afresh).

Dataset support for popular vision datasets such as COCO, Cityscapes, LVIS, PASCAL VOC, ADE20k.

Mask R-CNN is an instance segmentation model which extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. Most

importantly, Faster R-CNN was not designed for pixel-to-pixel alignment between network inputs and outputs. This is evident in how RoIPool, the de facto core operation for attending to instances, performs coarse spatial quantization for feature extraction. To fix the misalignment, Mask R-CNN utilises a simple, quantization-free layer, called RoIAlign, that faithfully preserves exact spatial locations.

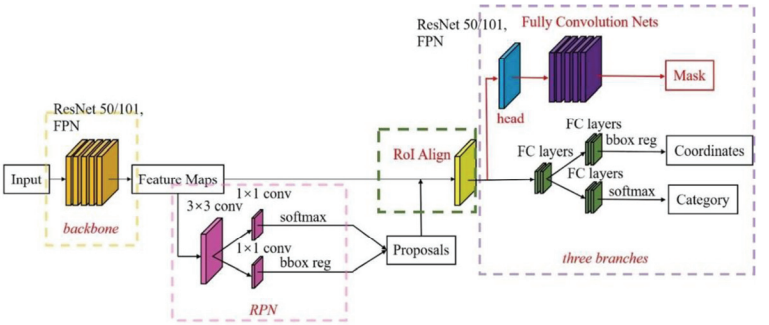

Figure 5. Mask R-CNN architecture

Loss function for each sampled RoI is:

$$L = L_{cls} + L_{box} + L_{mask} \tag{4}$$

3.3. Data preprocessing

To train the YoloV5 object detector, the data must be converted to the PascalVOC format [14].

Roboflow service [15] was used to perform this work. Roboflow service in the free version offers to upload a dataset of up to 10 thousand photos, split it for group partitioning, automatically compose partitioning files, select the necessary format (for which neural network the dataset is made), perform preprocessing (bring the data to the same size, select a filter), and also perform software augmentation of the dataset. We will consider the last function in more detail.

In order that there was no imbalance of classes, each class was represented evenly and weights were formed for each class equally the sample must be uniform. That is, for each of the required classes, an approximately equal number of photos is needed. Also, a large number of photos is needed to get good training results. For example, for the YoloV5 architecture, at least 1500 photos are needed for each class. In reality, it is a very difficult task to collect even frequent objects of high quality and diverse images. The solution to this problem is a process called augmentation - artificially increasing the size of the training data set by applying various filters or image processing algorithms. The simplest augmentation filter is to rotate the image by 90 degrees. For a neural network, this is already a different object to train, so even such simple filters are quite applicable.

As described earlier on Roboflow platform, there are tools for dataset augmentation, but there is a limitation on the multiplicity of augmentation (how many times to increase the size of the original dataset), and there is no possibility to normalize the sample by classes.

Also, with the help of built-in libraries of python and library for mathematical calculations numpy normalization of representativeness was made.

1. RESULTS & DISCUSSION

4.1. Experimental description

For the training of the YoloV5 object detector, a dataset was prepared based on the EMDS-7 work. From the original 2600 photos, the size of the dataset was increased to 26000 photos. Training the model on a computer with a discrete accelerator newer than 2019 lasts approximately 3-4 hours. Training for this work was performed on a machine learning server with eight Tesla A100-40 acceleration units specifically designed for AI, and took only 30 minutes. The achieved accuracy for 42 classes was 90% for data similar to EMDS-7, but the model performs poorly on third-party data (not least because of the green filter).

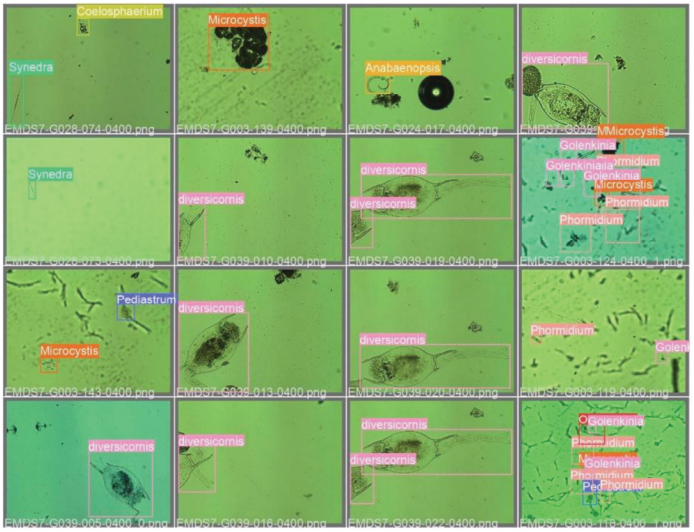

Figure 6. YoloV5 detector performance on EMDS7 data

To train the image segmenter, the same dataset preparation steps were applied as for training the YOLOv5 model, but now the training was done through the Detectron2 framework, the augmentation should consider that the markup format for segmentation is different, and the SinfNet dataset was taken as a basis. The result of the model based on SinfNet work is much better - due to more neutral background and image quality, this model worked on the found third-party photos of the same simple organisms, and no errors in labeling the contours of individual organisms were noticed.

The obtained segmentation model can already be integrated into a microbial detection program, but before that it is necessary to optimize the weights, a process called pruning. It is also worth testing the obtained models on manually collected and labeled data.

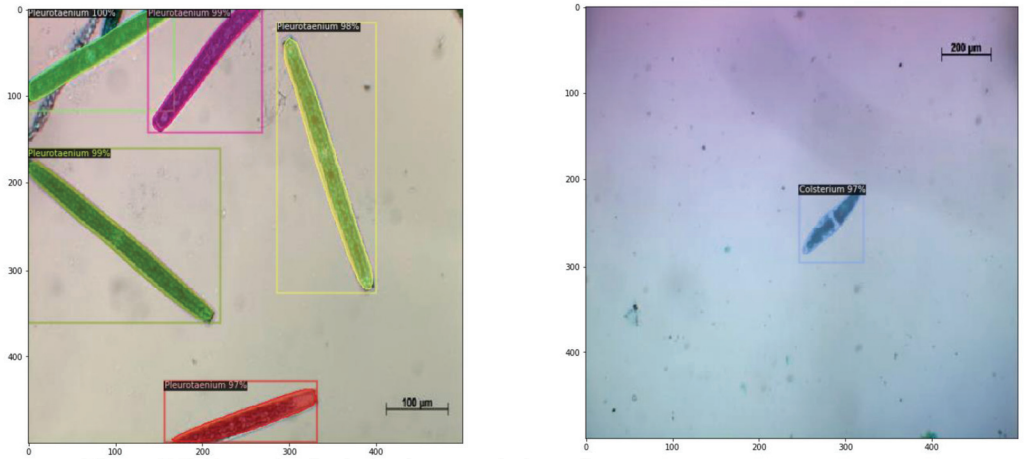

Figure 7-8. Example of a trained segmenter's work

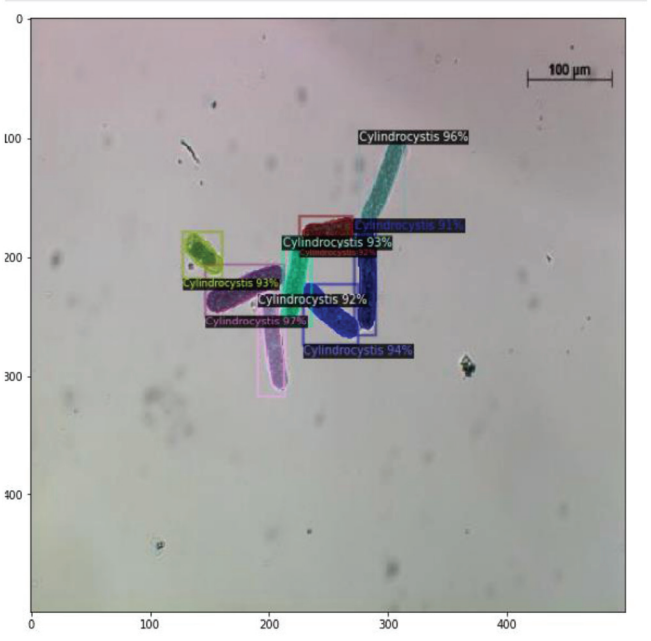

Figure 9. Example of a trained segmenter's work

1.2. Experimental description

To prove the performance of the trained protozoan microorganism detector model, it was necessary to collect our own preparations, take pictures, refine the model if necessary, and analyze the performance of the prototype.

The following equipment was selected as part of the research setup:

- Levenhuk D50L NG digital microscope, monocular giving a maximum magnification of 2400x and having artificial illumination of the table

- Digital camera d130 compatible with the eyepiece of the microscope for taking digital photos. The camera is amateur level, so the quality of photos does not correspond to the quality of laboratory cameras.
- Equipment for making preparations: slide glass, cover glass, Pasteur pipette.
- Biomaterial collection equipment: plastic bottles, forceps, knife.

Moskvoretsky Park, located in the North-West Administrative District of Moscow, was chosen as the place for biomaterial collection. There is a large number of ponds, swamps; there is a large coastline along which a large amount of living algae was found - places where the probability of finding microorganisms is much higher. As part of the collected biomaterial, 10 bottles of 100ml were taken, half of these bottles were collected along the shoreline of the Stroginsky Bay and the Moscow River, the other half in the marshes and puddles of the floodplain.

The methodology of creating the preparation was as follows:

- Take a small drop (approximately 3 ml) from the bottle with a pipette, preferably capturing a piece of algae in the drop. The algae can be pulverized beforehand for convenience.
- Pour the contents of the pipette in the shape of a crescent moon carefully in the center of the slide.
- Carefully place the coverslip on top of the slide without squeezing the contents too much.

The laboratory set-up and sample examination shown in Figure 10:

Figure 10. A microscope with a digital camera in action

As a result of the examination of the collected biomaterial, 11 minutes of video footage of live microorganisms were captured at a magnification of 400x. An example of the footage shown in Figure 11.

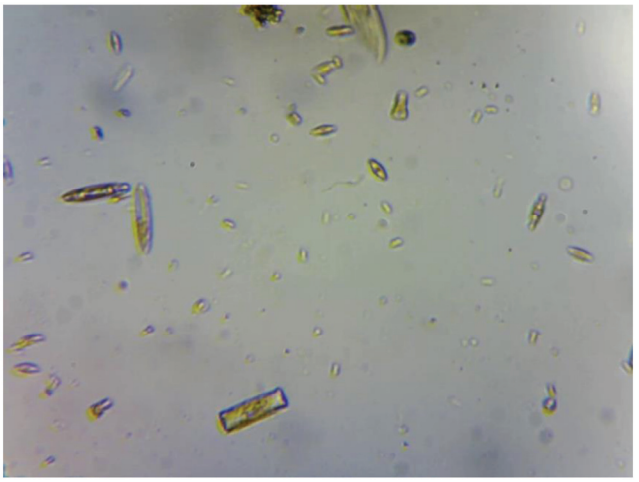

Figure 11. Example of footage

A large number of diatomic algae and a couple of protozoa like infusoria were filmed. All recorded videos were sliced using the ffmpeg video tool (moments with microorganisms were cut out), then split into individual frames.

Immediately afterwards, a small number of frames were fed to the previously trained YoloV5 model, but only the available classes in the Chinese dataset were triggered, namely one species of diatom algae. Because of this, it is necessary to retrain the model on the new microorganisms found. Retraining is the process of training a neural network by taking the original weights (already trained on some dataset) as input data and training it by building on the links already present. This improves the quality of recognition and adds new classes to the existing ones.

The previously prepared frames were uploaded to the Roboflow markup service, marked up, augmented with previously written scripts, and then we further trained the model on the new microorganisms found. The result of the model became much better, and now on each of the sliced videos the model began to correctly detect large microorganisms. An example of the detector performance on our own data is shown in Figure 12.

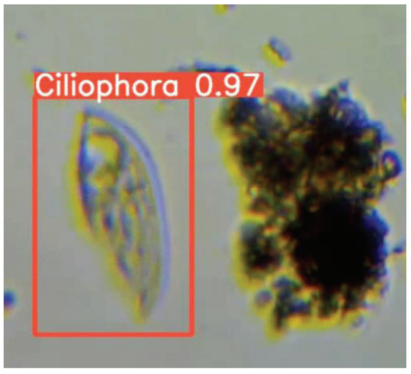

Figure 12. Example of detector operation on own data

Due to the large variety of diatom algae and the quality of the installed camera, it was not possible to obtain a qualitative detection of diatom algae. There were many false positives. Increasing the quality of the training material would greatly improve the situation. The results of the study are presented in the table 3.

Table 3: Experimental results

Model	Dataset	Accuracy
Yolov5	EMDS-7	90%
MaskRCNN	SnifNet	95%
Yolov5	EMDS-7 + custom dataset	93%

1.3. Discussion

According to the final results of the work, we can see that from the data available in the public domain, using standard algorithms for data augmentation on the scenarios of searching for microorganisms in the image neural networks give a good result for both detection and segmentation (subsequently applicable for biomass counting). But this result is possible only if there are similar data format and required microorganisms in the training sample. However, due to the fact that there is a huge number of microorganisms, and there is no single database for such purposes, when a microorganism, which is not in the database, appears when changing equipment, there is a need for additional training of the model.

5. CONCLUSION

Within the framework of this article the following tasks were performed: analysis of the subject area was carried out, general problems in solving this problem were deduced, solution options were proposed, a set of own data for training neural networks was collected on the basis of samples collected in protected areas, own standardization of data for microorganisms was developed, a prototype of neural network was created, which could solve the problem of classification of microorganisms in microscope analysis.

Since in solving many of the problems, a number of problems often encountered in solving such problems were discovered, the conclusion will describe ideas for further research in this area.

To solve the most important problem in the field of application of AI for solving biological problems, namely the problem of collecting a large amount of data, it is proposed to organize a universal database containing not only the characteristics of microorganisms, but also a set of photographs in a representative volume, necessary and sufficient for training of neural networks. Since this task will be very difficult to solve by the classical method: simply collecting a huge amount of information, it is proposed to introduce the possibility of intelligent data augmentation, namely the application of generative adversarial neural networks [16] for qualitative data augmentation that will be able to level out the insufficiency of data on specific equipment, on specific magnification, on specific type of microscopes used. Similar research on the application of GAN networks is already ongoing in medicine

[17], without clear successful results and global goals, but with the appearance of next-generation GAN networks, this idea can be realized and permanently solve the problem of data collection in biology.

As a direct extension of this research, it is possible to create an autonomous unit to remotely conduct sanitary surveys at water bodies or factories. The creation of a working prototype will be able to speed up the results of sanitary inspections, as well as to carry out tests immediately if necessary (for example, after the expiration of certification). Autonomization and digitalization of this process will reduce the costs of unnecessary logistics of preparations from a large number of locations, as well as to free qualified personnel from routine work, for more complex and scientific tasks.

5. REFERENCES

Chu X., Liu X., 2022, *Detection of Cells and Microbes in Microscopic Field Based on Improved YOLOv5.* 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA).

Tasnim A., Md Wahid F., Md Hasan J., 2019, *Combining Deep Convolutional Neural Network with Support Vector Machine to Classify Microscopic Bacteria Images*.1-5.10.1109/ECACE.2019.8679397.

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*. Sensors and Actuators B: Chemical; 309:127789.

Mhathesh T., Onesimu A., Sagayam M., Henesey L., 2021, *A 3D Convolutional neural network for bacterial image classification,* 10.1007/978-981-15-5285-4_42

Zieliński B., Plichta A., Misztal K., Spurek P., Brzychczy-Włoch M., Ochońska D., 2017, *Deep learning approach to bacterial colony classification*. PLOS ONE. 12. e0184554. 10.1371/journal.pone.0184554.

Bodas A., Shubhankar H., Rujuta S., Joglekar A., Neeta S., 2022, *A Computer Vision Model for Detection of Water Pollutants Using Deep Learning Frameworks*. Lecture notes in networks and systems (Online); 543–53.

Yuan F., Huang Y., Chen X., Cheng E., 2018, *A Biological Sensor System Using Computer Vision for Water Quality Monitoring*. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2876336.

Irani T., & Amiri H., Azadi S., Bayat M., Deyhim H., 2022, *Use of a convolution neural network for the classification of E. Coli and V. Cholara bacteria in wastewater*. Environmental Research and Technology. 5. 10.35208/ert.969400.

Yang H., Li C., Zhao X., Cai B., Zhang J., Ma P., Zhao P., Chen A., Jiang T., Sun H., Teng Y., Qi S., Huang X., Grzegorzek M., 2023, *EMDS-7: Environmental microorganism image dataset seventh version for multiple object detection evaluation*. Frontiers in Microbiology. 14. 1084312. 10.3389/fmicb.2023.1084312.

Sabban S., Alotebi A., Khalifah Z., Alafif T., 2023, SinfNet: Microorganism image classifier [Internet]. GitHub. 2023 [cited 2024 Aug 6]. Available from: https://github.com/sarisabban/SinfNet

Spahn C., Laine R., Pereira P., Gómez de Mariscal E., Chamier L., Conduit M., Pinho M., Holden S., Jacquemet G., Heilemann M., Henriques R., 2021, *DeepBacs: Bacterial image analysis using open-source deep learning approaches*. 10.1101/2021.11.03.467152.

World Data Centre for Microorganisms [Internet]. www.wdcm.org. [cited 2024 Aug 6]. Available from: https://www.wdcm.org

Gallardo R. DIBaS-Dataset [Internet]. GitHub. 2023. Available from: https://github.com/gallardorafael/DIBaS-Dataset

Everingham M., Van Gool L., Williams C., Winn J., Zisserman A., 2010, *The Pascal Visual Object Classes (VOC) challenge*. International Journal of Computer Vision. 88. 303-338. 10.1007/s11263-009-0275-4.

Roboflow: Give your software the power to see objects in images and video [Internet]. roboflow.com. Available from: https://roboflow.com

Creswell A., White T., Dumoulin V., Arulkumaran K., Sengupta B., Bharath A., 2017, *Generative Adversarial Networks: An Overview*. IEEE Signal Processing Magazine. 35. 10.1109/MSP.2017.2765202.

Lorencin I., Baressi Šegota S., Anđelić N., Mrzljak V., Ćabov T., Španjol J., Car Z., 2021, *On Urinary Bladder Cancer Diagnosis: Utilization of Deep Convolutional Generative Adversarial Networks for Data Augmentation.* Biology. 10. 10.3390/biology10030175.