# Improving generalization in classification of novel bacterial strains: a multi-headed ResNet approach for microscopic image classification

*V.O. Yachnaya* [1,2], *M.A. Mikhalkova* [1], *R.O. Malashin* [1,2], *V.R. Lutsiv* [2], *L.A. Kraeva* [3,4],
*G.N. Khamdulayeva* [3], *V.E. Nazarov* [5], *V.P. Chelibanov* [6]

[1] *Pavlov Institute of Physiology, Russian Academy of Sciences,*
*199034, Saint Petersburg, Russia, Naberezhnaya Makarova 6;*
[2] *Saint Petersburg State University of Aerospace Instrumentation,*
*190000, Saint Petersburg, Russia, Bolshaya Morskaya 67;*
[3] *Saint Petersburg Pasteur Institute, 197101, Saint Petersburg, Russia, Mira street 14;*
[4] *Military and Medical Academy named after S.M. Kirov, 194044, St. Petersburg, Russia;*
[5] *North-Western State Medical University named after I.I. Mechnikov,*
*191015, Saint Petersburg, Russia, Kirochnaya street 41;*
[6] *ITMO University, 197101, Saint Petersburg, Russia, Kronverksky prospekt 49*

## Abstract

The purpose of this work is to design a system for microscopic bacterial images classification that can be generalized to new data. In the course of work, a dataset containing 23 bacterial species was collected. We use a strain-wise method for dividing the dataset into training and test sets. Such splitting (in contrast to random division) allows evaluating the performance of classifiers on new strains in the case of intra-species visual variability of bacteria. We propose a "Multi-headed" ResNet (ResNet-MH) for the analysis of microscopic images of bacterial colonies. This approach forces the neural network to analyze features of different resolutions, such as the shape of individual bacterial cells and the shape and number of bacterial clusters during training. Our network achieves the 41.6 % accuracy species-wise and 64.06 % accuracy genera-wise. The proposed method of dataset splitting guarantees generalization to new unseen strains, whereas random splitting into training and test sets leads to overfitting of the system (accuracy is over 90 %). For the 10 visually strain-wise stable species, the accuracy of the proposed system reaches 83.6 % species-wise.

*Keywords*: bacteria classification, image classification, deep neural network, dataset splitting, multi-head model, microscopic images.

## Introduction

Bacterial genera and species recognition is essential for medicine, veterinary science, biochemistry, food industry and agriculture. Some bacteria species are the main cause of various diseases, so the accurate and rapid bacteria identification plays an important role in diagnosing and prescribing appropriate therapy.

Microbiologists specify the bacterial biochemical and morphological features for bacteria genera and species identification. For example, the appearance of colonies and microscopy of preparations (slides) from a pure culture of grown bacteria allow to estimate to which large taxon microorganism belongs. The bacteria then are to be identified by, for example, biochemical testing. This task requires an expert assessment by microbiologists with relevant expertise and experience, as well as special equipment, growth media and reagents and a strict compliance with cultivation procedures.

The purpose of this study is to develop a method of automatic bacterial species classification based on microscopic images of their colonies.

The contribution of the authors of this work is three-fold. Firstly, a new dataset of microscopic images of bacterial colonies was collected. Secondly, we have developed a neural network classifier and a data preprocessing method to recognize bacterial species. Thirdly, we show that bacterial strains have strong variability within one species, so it is required to form a test set correctly. The conducted experiments show that modern machine learning methods allow generalizing data beyond the strains contained in the training set even with limited data available, and the proposed neural network training method reduces the effect of overfitting and thereby improves the accuracy of the system. However, the dataset splitting, which takes into account the visual variability of strains within the genera, reduces the recognition accuracy, but increases the ability of the system to generalize on new data.

A main feature of the proposed neural network architecture is the multi-headed training, which forces to form features to analyze the microstructure (the shape of bacterial cells) and the macrostructure of the image (the shape and number of colonies).

## 1. Related works

Various means and methods are used in microbiology to study microorganisms, in particular bacteria. The tasks are to detect and identify bacterial cells, determine their properties, morphology, describe life cycle etc. The main methods are spectrometry and image analysis. The first group of methods includes, for example, analysis of mass spectrometry data as demonstrated in [1] or Raman spectrometry data in [2]. Methods of the second group fall into one of two categories: analysis of microscopic images of bacterial colonies and analysis of images of macroscopic studies (hereinafter referred to as macroscopic images).

The examples of works that process macroscopic images are [3 – 6]. The main tasks in this case are the detection of bacterial colonies as in [6], determination of the morphological properties of colonies like in [4, 6], tracking colonies growth and motility in time as in [5, 6].

The paper [7] contributes to the microscopic bacterial image classification via texture analysis. The authors conclude that important species-specific information is contained not only in the shape of individual bacteria, but also in in the shape of the colonies and their density. Digital Images of Bacteria Species (DIBaS) dataset [8] collected for experiments contains 660 images for 33 bacterial species of 20 different genera; slides were prepared using Gram staining technique.

DIBaS dataset was a source for developing several bacteria recognition methods: Support Vector Machine (SVM) based as in [7], the histogram equalization and bag-of-words feature extraction as in [9], a modification of the VGG-16 model as in [10]. In [11] authors augmented the DIBaS dataset and designed a compact classification model of 10 species with accuracy of more than 96 %. Despite high reported accuracies of the classifiers, we admit that DIBaS does not contain information on bacterial strains. As we show, ignoring this information does not allow proper estimation of the strain-wise generalization abilities (at least for non-prepared bacteria).

Images of Gram-stained bacterial colonies of three different species are also analyzed in [12].

A number of publications are focused on recorded live images and time-lapse videos of bacteria [13 – 22]. The paper [13] implements a method to count bacterial cells in the image fragment by means of classification. The works [15, 17 – 19, 21] are focused on the analysis of the microcolonies growth dynamics, which also involves counting the bacterial cells. For example, in [17] and [19] neural network methods for bacterial cells segmentation and subsequent estimation of their number are implemented. In [18], classical computer vision methods (such as Contrast-limited adaptive histogram equalization (CLAHE) and the Otsu's method) are used to segment bacterial colonies and determine their characteristics. In [21], ImageProPlus medical image tool estimates the heterogeneity in the growth behavior of single cells. The work [15] examines the applicability of neural networks for various tasks of bright-field and fluorescent images processing, including bacterial cell segmentation and object detection. For segmentation purposes they apply U-Net as demonstrated in [23], Content-Aware image Restoration (CARE) [24], pix2pix [25], StarDist [26], SplineDist [27]. The object detection methods allow distinguishing growth stages: the YOLOv2 (You Only Look Once) network [28] is used to discriminate between rod-shaped cells, dividing cells, and microcolonies. The Python package from [16] suits for purposes of bacterial fluorescence microscopy data and offers several visualization methods.

The paper [14] suggests classification methods to recognize 5 bacterial species of 5 different genera. The authors compare Long Short-Term Memory (LSTM) as in [29], Residual Neural Network (ResNet-18) as in [22] and One-dimensional Convolutional Neural Network (1D-CNN), which achieve recognition accuracy of 92.2 %, 93.8 % and 96.2 % respectively. It should be noted that the cells do not form bacterial clusters (microcolonies) in the analyzed images, but are located at a distance from each other.

The classification of bacteria is also the subject of [20]. However, in this case, the images are differentiated into four classes depending on the cell wall constituents and shape. The authors omitted many morphotypes from consideration due to lack of data. The convolutional neural network developed by the authors achieves 94.9 % accuracy for such classification.

Methodology used in a particular laboratory specifies different parameters to capture bacterial colonies: duration of the cultivation process, scale of the images, etc. These parameters should be taken into account and researchers should be able to cultivate and sample microorganisms appropriately in case of using off-the-shelf solutions. At the same time, dataset gathered from various online sources (for example, [30]) may contain heterogeneous data, various elements of which were collected under certain specific conditions. In this case, it is important to consider that bacterial species represented by only one preparing and recording method will require researchers to follow a similar data preparation process, which is not possible in case of examining samples of a class unknown in advance.

Note that in the considered papers [7, 14, 20], classifiers are tested on randomly selected (in terms of strains) from the dataset's examples. This approach makes it possible to achieve high classification results, but does not take into account the intraspecies variability of bacterial strains, in case samples of the same strain reach both the training and test sets. This can significantly affect the recognition accuracy of images obtained from new cultures. In our research we take into account the intraspecies visual variability of bacterial clusters depending on the bacterial strain.

## 2. Material and methods
### 2.1. Dataset of bacterial colonies microscopic images

The microbiological part of this study was carried out in the Laboratory of Medical Bacteriology of the Saint-Petersburg Pasteur Institute and consisted of the following:

1. All bacterial strains were studied and identified using standard methods: study of colony morphology, microscopy of stained preparations from grown microorganisms, biochemical identification using biochemical testing kits: MIKROLATEST® test kits for bacteria identification (ERBA LACHEMA), Microbact (BioVitrum) kit for biochemical identification, API® strips (BioMerieux). All obtained isolates were additionally identified using the mass spectrometry technique on a MicroflexTM LT MALDI-TOF mass spectrometer (Bruker Daltonics, Germany) using the Flex Control program (the device was operated in a linear positive mode with the necessary parameters described in the instructions for the device).

2. Strains of identified bacteria were grown for 3 hours on meat-peptone agar with the addition of 10 % bovine serum at a temperature of +37°C, the density of the bacterial mass during inoculation was 1•106 CFU/ml (CFU – colony-forming units).

3. Grown microcolonies registration was carried out using an Axio Scope A1 microscope (manufactured by Zeiss) with magnification up to 400x and an AxioCamHRc Rev3 professional stationary digital camera using a special N-achroplan objective to work with a three-dimensional image. In this case, the identificator of each microorganism strain was registered. For every strain several images without overlapping field of view were captured.

This technique suits prompt diagnosis, as the growth time is low enough to observe the bacterial colonies and the common equipment is used in [30].

We collected 4618 microscopic images of 23 bacterial species of 11 different genera. Each species is represented by at least two different strains. Strain is a culture within a biological species, having special physiological and biochemical characteristics. By default, one assumes that different patients possess different strains of bacteria, which may belong to the same species. Therefore, generalization over different strains has the great importance in practical applications. We consider task of recognition of 23 image classes related to 23 bacteria species.

The images contained in the dataset are colored JPG format images with a 1388×1040 resolution. It should be noted that the color of the images depends on settings of the camera, lighting and the nutrient medium for the microorganisms and therefore does not indicate bacterial species or genera. Detailed description of the dataset, such as number of strains and images for each species, is presented in Tab. 1.

The microorganisms presented in the dataset are of different morphotypes (rods and cocci) and different types of colony structure, which characterize the bacterial species.

*Tab. 1. Dataset description*

| №g | Genera | №s | Species | Number of strains | Number of images |
|---|---|---|---|---|---|
| 1 | Acinetobacter | 1 | baumannii | 2 | 127 |
| | | 2 | pittii | 3 | 122 |
| | | 3 | seifertii | 2 | 41 |
| | | 4 | ursingii | 2 | 50 |
| 2 | Candida | 5 | albicans | 2 | 130 |
| 3 | Citrobacter | 6 | braakii | 2 | 40 |
| | | 7 | freundii | 2 | 50 |
| 4 | Enterobacter | 8 | cloacae | 7 | 399 |
| 5 | Enterococcus | 9 | faecalis | 9 | 375 |
| | | 10 | faecium | 7 | 241 |
| 6 | Escherichia | 11 | coli | 8 | 476 |
| 7 | Klebsiella | 12 | aerogenes | 2 | 61 |
| | | 13 | pneumoniae | 6 | 337 |
| | | 14 | variicola | 2 | 33 |
| 8 | Proteus | 15 | mirabilis | 5 | 276 |
| 9 | Pseudomonas | 16 | aeruginosa | 10 | 365 |
| 10 | Staphylococcus | 17 | aureus | 8 | 320 |
| | | 18 | epidermidis | 5 | 249 |
| | | 19 | haemolyticus | 2 | 199 |
| | | 20 | hominis | 2 | 213 |
| 11 | Streptococcus | 21 | agalactiae | 9 | 315 |
| | | 22 | anginosus | 2 | 110 |
| | | 23 | pneumoniae | 3 | 89 |

**Dataset splitting.** The dataset splitting was carried out strain-wise. Thus, for each of the species, all examples of a strain go either into the training set or into the test set. Since each strain is unique and may have different properties, such splitting technique allows to obtain a more reliable estimate on new strains, in contrast to random image splitting of the entire dataset into training and test sets. At the same time, we divided training set into training and validation ones randomly, regardless of the strain.

Visual variability of strains within many bacterial species includes different number of clusters of cells, differences in shape and size to the same time of taking images. Example of visually varying strains of the same species is shown in Fig. 1a. The problem of variability within a species depending on a strain is also noted in [4, 6].

As seen from Fig. 1a, the variability of strains is so high in some cases that it makes it impossible to recognize another strain only from a single strain sample in the train set. However, we show that many bacterial species have less variability and can be successfully classified. For example, Fig. 1b presents example of class for which the visual features of strains are stable.

We use a hold-out cross-validation for dataset division for further experiments. A small number of strains in the database leads to the fact that the test set contains only one strain, which, in turn, leads to a relatively low confidence in estimates, but allows maintaining the experiment's repeatability. This is important for model development, since different neural network weight initializa-

tion also affects the resulting recognition accuracy. We don't use k-fold cross validation due to the limited computational resources, as well as due to the fact that different classes are represented by different numbers of strains (which leads to multiple possible partitions).
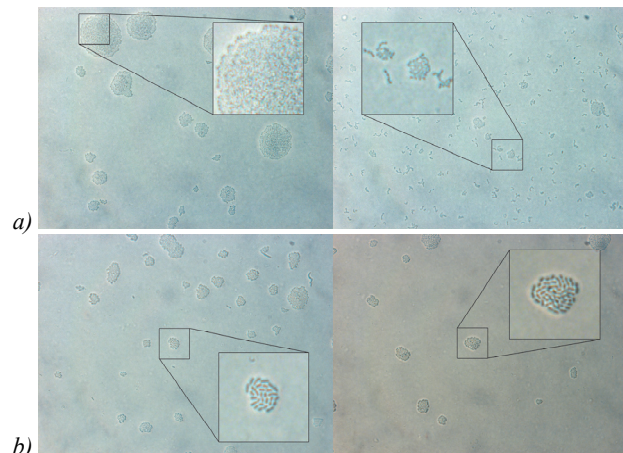


*Fig. 1. Examples of images contained in dataset: (a) Enterococcus faecalis strains (visually variable); (b) Klebsiella aerogenes strains (visually stable)*

For each class, the test set consists of one strain. As a result of the described dataset division, the number of images in the training set varies in the range from 20 to 407 (average – 172), in the test set – in the range from 8 to 99 (average – 29).

**Dataset augmentation.** The classes presented in the dataset are unbalanced in the number of images. In particular, the prior probability of the bacterial species varies, which leads to the fact that some classes are much more common than others. To avoid usage of class balance techniques during training we applied the following image augmentation methods (up to 400 images per class): rotation by 0 to 360 degrees, vertical or horizontal flip, vertical or horizontal shift by up to 10 %.

All the described transformations do not corrupt the features that characterize certain bacterial classes.

Moreover, on-the-fly augmentation was additionally applied during the training phase. It includes color transformations of the training set images. We find out that additional geometric augmentation on-the-fly doesn't improve generalization abilities on the test set.

In case of the strain-wise dataset splitting, the test set is also unbalanced in terms of the number of images. It can lead to the incorrect interpretation of the classifier accuracy on the test set. Hence, a class-wise recognition accuracy is also considered for models that achieve the best results.

### *2.2. ResNet-MH: forced analysis of different hierarchical features in the ResNet*

Some species strains contain a large number of colonies over the entire image area, others have very few of them. This can lead to the overfitting: neural network will use the number of colonies as a key factor in the process of weight adjustment, ignoring meaningful aspects.

More than that, an important feature that distinguishes bacterial species is not only the shape of the colony, but also the shape of the individual microorganism within the colony (Fig. 1).

For the simultaneous analysis of both the shape of colonies and individual microorganism, and even the mutual position of individual colonies, one can design a system that finds the most informative region of image, and then classifies it using a model that is trained on small fragments.

Such approach has a number of drawbacks including independent tuning of the mutually related modules (detection and classification) as demonstrated in [31].

For end-to-end detection and segmentation algorithms additional labeling is required, which in turn requires the involvement of experts, as such labeling is not obvious in some cases.

As alternative, we propose a ResNet-18 modification that provides simultaneous feature learning on different resolutions. In this section, we describe the designed solution, and in the section 4 we present the results of experiments and discus the modifications made. The structure of the proposed multi-headed ResNet (ResNet-MH) architecture is shown in Fig. 2.
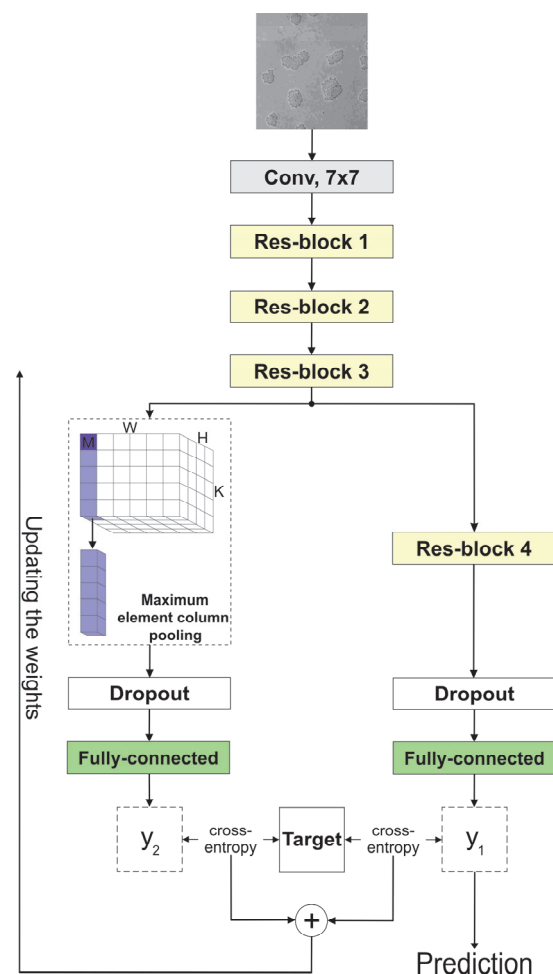


*Fig. 2. Scheme of the ResNet-MH architecture, where $y_1$ is the main output and $y_2$ is the additional output of the network*

It is different from classic ResNet in two aspects: it has additional outputs with additional dropout layers.

The purpose of all outputs $y_i$ is to predict the true label. In [32] similar additional classifiers allowed training deeper networks without residual blocks (Res-blocks). In this work, we have a different motivation: as soon as receptive field of neurons in the column is restricted, the additional output forces the network to analyze the images at a lower resolution, without the possibility of generalizing the data from the entire image. Thus, not the entire image is processed at the additional output, but a smaller fragment.

The receptive field of a neuron in a random column can contain an empty background without any colony. For the correct classification of a fragment, it is necessary to rely on a neuron that "observes" exactly the colony and gives the maximum response to it. For this purpose, a modified operation of maximum pooling is applied – maximum element column pooling. The Equations 1 and 2 describe maximum element column pooling (denoted as "MECP"):

$$r,c,d = \operatorname{argmax}_{i,j,k}\left(X_{i,j,k}\right), \tag{1}$$

$$\operatorname{MECP}(X): \mathbf{X} \to \mathbf{X}_{r,c}, \tag{2}$$

where $r$, $c$ and $d$ are indexes of the row, column and depth respectively of the input tensor $\mathbf{X}$.

It is assumed that the maximum response occurs in the informative region. We investigated other ways to determine the informative region, but they appeared to be less effective (we discuss it in the section 4).

The training of the developed neural network was carried out by optimizing the loss function L:

$$L = L_4 + \sum_{i=1}^{3}\lambda_i L_i, \tag{3}$$

where $L_i$ is the cross-entropy loss produced after $i$-th Res-block, $\lambda_i$ is hyperparameter.

First output produces the final prediction, as it analyses all the information, while the other outputs improve the features on which the main classifier relies.

It was experimentally concluded that for the given dataset the architecture shown in Fig. 2 (the second output is added after the third Res-block) achieves the best results. In our experiments the best result was obtained with two-headed ResNet ($\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 0.3$).

Dropout layers before the fully connected layers stabilize training. The results of these experiments are described in the next section.

The proposed modifications are not specific to ResNet architectures and can be applied in a similar way to almost any convolutional neural network.

Noteworthy, while the multi-headed approach is a common feature in ensemble learning for tasks such as distillation [33] or domain generalization [34], where backbone-network performs feature extraction and the heads recognize the domain's specificity, our approach varies. In these conventional scenarios, introducing additional heads is a general methodology. In contrast, we add heads due to the homogenous images' microstructure of our domain [35] and small training set. In our approach, it is essential to attach extra heads to features corresponding to smaller receptive fields, while conventionaly all the heads are fed with the features of the same levels.

### 3. Experimental results

For the experiments wherever otherwise stated, we used the following training hyperparameters:
- the model is pretrained on the ImageNet dataset [36];
- number of epochs – 30;
- input image size – 1024×1024 pixels;
- color mode of the input image – grayscale;
- batch size – 23;
- optimization algorithm – Adaptive Moment Estimation as in [37];
- learning rate – 1e-4;
- weight decay – 1e-5.

We repeat every experiment with different weight initialization from 3 to 8 times to obtain statistically valid results.

We control accuracy on the test set after each epoch during the training process and report three metrics for comparative analysis:
1. mean maximum accuracy (MMA) for the test set (for each run selecting best epoch and average across runs);
2. maximum accuracy on the test set for all training epochs and all runs;
3. average Top-1 accuracy for training, validation and test sets (averaging across runs and epochs from the epoch when the training accuracy plateaus).

The training process was performed on NVIDIA TITAN X (Pascal) and NVIDIA GeForce GTX 1080 GPUs.

We conducted a series of experiments aimed at determining the most suitable classifier architecture and some data parameters for solving our task. First, we selected the classifier architecture from the following ones: MobileNetV2 as in [38], ResNet-18 and ResNet-50 [22], EfficientNet-B7 [39]. Secondly, we evaluated color mode and image resolution effect on classification accuracy. As a result of the examination, our further work is carried out with the ResNet-18 model, which processes grayscale images with a resolution of 1024×1024 pixels. The experiments are described in more detail in **Appendix A**.

### *3.1. Multi-headed ResNet*

We studied the following important aspects of the designed ResNet-MH architecture: the method for finding the most informative region, hyperparameter vector λ.

**Dropout.** In the course of the described above experiments we noticed that the model converges at the early training epochs. It was suggested that adding dropout layer before the fully connected layer can slow down the training and reduce overfitting as demonstrated [40]. A similar modification is used, for example, in [41], where the authors also built a classifier for microscopic studies.

We observed better training dynamics in ResNet-18 with dropout compared to original version, and MMA slightly increased (by 0.2 %).

**Loss configuration.** We studied opportunity to use additional heads after 2nd and 3rd Res-blocks (we considered that receptive field after the 1st block is too small and therefore assumed $\lambda_1 = 0$ in Equation 3 of the loss function). The receptive field of a neuron is smaller in the former case ($99 \times 99$ vs $355 \times 355$), and thus cannot contain more than one colony (i.e., excessive information). However, the modification with the additional output set up after the third Res-block reaches a better classification accuracy (Tab. 2).

*Tab. 2. Classification accuracy by resnet-18 and its modifications*

| Model | Pretraining dataset | Acc. (test), % | MMA (test), % | Max.acc. (test), % |
|---|---|---|---|---|
| ResNet-18 | ImageNet | 34.47 | 36.9 | 41.2 |
| ResNet-18 | CytoImageNet | 30.56 | 31.92 | 33.39 |
| ResNet-18 with dropout | ImageNet | 35.4 | 37.1 | 38.6 |
| ResNet-MH ($\lambda_1 = \lambda_3 = 0$, $\lambda_2 = 0.3$) | ImageNet | 34.99 | 37.46 | 39.4 |
| ResNet-MH($\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 0.3$) | ImageNet | **36.4** | **38.4** | **41.6** |

The developed ResNet-MH outperforms the best considered model – ResNet-18 with dropout. The results of training the ResNet-18 and its modifications are shown in Tab. 2; the values are averaged over the results of at least 8 training runs for each of the models.

**Weights transfer.** We considered the effect of transfer learning. All the models described earlier were pretrained on the ImageNet dataset. We tried to substitute ImageNet with dataset of microscopic images CytoImageNetas in [42] that contains 1000 images for each of 894 classes. The results of ResNet-18 model pretrained on this dataset is shown in Tab. 2. Though CytoImageNet covers domain of microscopic images we didn't observe any benefit from pretraining on it.

**Alternative methods to find the most informative region.** As described in the previous section, maximum element column pooling selects the informative region of the fragment which is considered at the additional output of the designed architecture (the second output was added after the third Res-block). We carried out experiments with some other possible methods.

Firstly, we compare "signed" maximum (Equations 1 and 2) and absolute maximum for column selection. The Equations 4 and 5 describe absolute maximum element column pooling (denoted as "AMECP").

$$r, c, d = \mathrm{argmax}_{i,j,k} \left| X_{i,j,k} \right|, \tag{4}$$

$$\mathrm{AMECP}(X): \mathbf{X} \to \mathbf{X}_{r,c}, \tag{5}$$

where $r$, $c$ and $d$ are indexes of the row, column and depth respectively of the input tensor $\mathbf{X}$.

The other approach involves attaching additional fully connected layer to each column; layer had 23 output neurons and was shared across all the columns. This layer was optimized to minimize cross entropy loss during training. Minimal entropy of this additional predictor served as a signal of "informativeness" of the column and underlying image region.

Tab. 3 shows the results of these experiments. Here, "Entropy" denotes minimum entropy pooling.

*Tab. 3. Classification accuracy depending on the most informative region choosing method (averaging over three training runs)*

| Method | Acc. (val), % | Acc. (test), % | MMA (test), % |
|---|---|---|---|
| MECP (ResNet- MH) | 97.7 | 36.4 | 38.4 |
| AMECP | 97.3 | 32.4 | 34.3 |
| Entropy | 95.7 | 34.9 | 37.13 |

As can be seen from Tab. 3, simplest maximum element column pooling provides the highest accuracy, we use this method in further experiments with the Multi-headed ResNet.

### 4. Discussion

Fig. 3*a* shows the variance and average classification accuracy for the ResNet-18 architecture and its modifications over 8 different training runs with different random seeds. Fig. 3*b* shows the MMA. As can be seen from Fig. 3, the ResNet-18 has the highest accuracy variance, while adding a dropout layer reduces it. Fig. 3*b* also shows that the MMA for the designed ResNet-MH (Fig. 2) is about $1 - 1.5$ % greater than for other architectures considered.
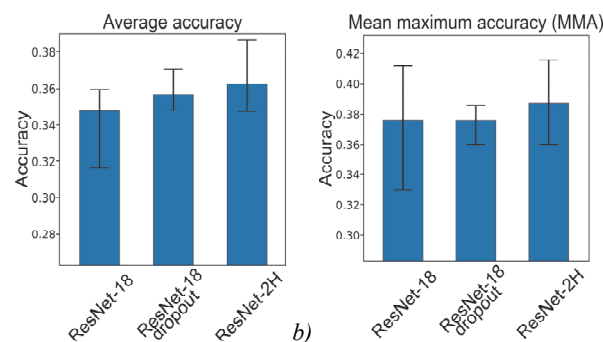


*Fig. 3. Classification accuracy of different architectures comparison*

A confusion matrix of the ResNet-MH is shown in Fig. 4 (the class numbers correspond to the species numbers assigned in Tab. 1). We use the model that gives the best Top-1 accuracy on the test set (41.6 %) as an example.

As can be seen from Fig. 4, 10 classes are completely misclassified. However, some of them are identified correctly within the genera (for example, class 9 *Enterococcus faecalis*). It indicates the visual similarity of bacterial species within the same genera. Generalizing classes up to genera increases classification accuracy of this model to 64.06 %.

Classifier predictions

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 | 0 | 20 | 40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 30 | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 40 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 70 | 0 | 5 | 0 | 0 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0 | 38 | 0 | 0 | 0 | 0 | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 63 | 0 | 0 | 12 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 90 | 5 | 0 | 0 | 5 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 50 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 60 | 30 | 10 | 0 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 25 | 20 | 0 | 5 | 0 | 0 | 0 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 85 | 0 | 0 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 0 | 3 | 0 | 0 | 8 | 0 | 18 | 0 | 20 | 0 | 0 | 20 | 25 | 0 |

*Fig. 4. ResNet-MH confusion matrix. Target classes are located along the vertical axis, classifier predictions are located along the horizontal axis, the recognition accuracy of the target class (%) is inside the cells. Bold lines indicate the "boundaries" of the bacterial genera*

However, there are the cases of incorrect classification even among the genera, which results from the visual similarity of their images. For example, Fig. 5*a* shows a test image of the *Acinetobacter pittii* class, which was identified as of the *Enterococcus faecium* class, and Fig. 5*b* shows a training example of the *Enterococcus faecium* class for comparison.

The visual strains variability within species affects the recognition accuracy. We expect that scaling dataset may significantly increase classification accuracy, but we consider that under specified shooting conditions some species may remain visually undistinguishable.
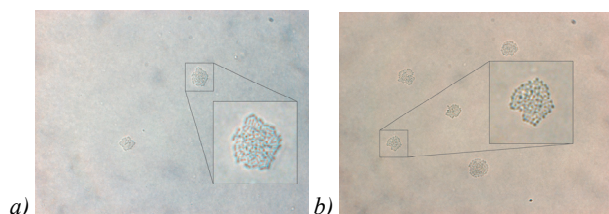


*a)* *b)*

*Fig. 5. Visual similarity of images of different genera: (a) Example of Acinetobacter pittii class; (b)Example of Enterococcus faecium class*

After excluding from consideration the classes for which the recognition accuracy is less than 50 %, the recognition accuracy of the remaining 10 classes by the developed model reaches 83.6 % species-wise and 89.7 % genera-wise.

## Conclusion

In this paper we design a bacterial classifier capable of generalization on new data, namely on new bacterial strains. The developed ResNet-MH architecture for the bacterial microscopic images analysis uses the additional output (second head), which makes it possible to analyze the micro- and macrostructure of images, extracting fea-tures of both bacterial cells and bacterial clusters. The designed maximum element column pooling ensures that the fragment of image with an empty background is ignored.

To evaluate the classifier, a dataset of microscopic images of 23 bacterial species belonging to 11 genera was collected. Due to the high intraspecies visual variability of bacterial cells, we perform the strain-wise dataset split-ting, while the random splitting has proved to lead to the neural network overfitting and low generalizing ability for new strains. Although in contrast to other published bacteria classification algorithms our recognition results are lower, we consider the case of complex bacterial strains visual variability and force the system to general-ize on the new data.

Our ResNet-MH achieves 41.6 % recognition accura-cy species-wise and 64.06 % genera-wise, which is higher than other classification methods tested. Moreover, the recognition accuracy reaches 83.6 % for the 10 most vis-ually stable species.

In the future, we are planning to expand the dataset and improve the developed method, by additional use of image segmentation methods.

## References

[1] De Bruyne K, Slabbinck B, Waegeman W, Vauterin P, De Baets B, Vandamme P. Bacterial species identification from MALDI-TOF mass spectra through data analysis and machine learning. Syst Appl Microbiol 2011; 34(1): 20-29. DOI: 10.1016/j.syapm.2010.11.003.

[2] Ho CS, et al. Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning. Nat Com-mun 2019; 10(1): 1-33. DOI: 10.1038/s41467-019-12898-9.

[3] Sajedi H, Mohammadipanah F, Pashaei A. Image-processing based taxonomy analysis of bacterial macro-morphology using machine-learning models. Multimed Tools Appl 2020; 79(43): 32711-32730. DOI: 10.1007/s11042-020-09284-9.

[4] García-Soriano DA, Andersen FD, Nygaard JV, Tørring T. ColFeatures: Automated data extraction and classification of bacterial colonies. bioRxiv Preprint. 2021. Source: <https://www.biorxiv.org/content/10.1101/2021.05.27.445853v1>. DOI: 10.1101/2021.05.27.445853.

[5] Doshi A, et al. A deep learning pipeline for segmentation of Proteus Mirabilis Colony Patterns. IEEE 19th Int Symp on Biomedical Imaging (ISBI) 2022: 1-5. DOI: 10.1109/ISBI52829.2022.9761643.

[6] Bär J, Boumasmoud M, Kouyos RD, Zinkernagel AS, Vulin C. Efficient microbial colony growth dynamics quantification with ColTapp, an automated image analysis application. Sci Rep 2020; 10(1): 1-15. DOI: 10.1038/s41598-020-72979-4.

[7] Zieliński B, Plichta A, Misztal K, Spurek P, Brzychczy-Włoch M, Ochońska D. Deep learning approach to bacterial colony classification. PLoS ONE 2017; 12(9): e0184554. DOI: 10.1371/journal.pone.0184554.

[8] DIBaS dataset. 2017. Source: <http://misztal.edu.pl/software/databases/dibas/>.

[9] Mohamed BA, Afify HM. Automated classification of bacterial images extracted from digital microscope via bag of words model. 9th Cairo Int Biomedical Engineering Conf (CIBEC) 2018: 86-89. DOI: 10.1109/CIBEC.2018.8641799.

[10] Patel S. Bacterial colony classification using atrous convolution with transfer learning. Ann Rom Soc Cell Biol 2021; 25(4): 1428-1441. Source: <https://www.annalsofrscb.ro/index.php/journal/article/view/2650>.

[11] Mai D-T, Ishibashi K. Small-scale depthwise separable convolutional neural networks for bacteria classification. Electronics 2021; 10(23): 3005. DOI: 10.3390/electronics10233005.

[12] Song Y, et al. Segmentation, splitting, and classification of overlapping bacteria in microscope images for automatic bacterial vaginosis diagnosis. IEEE J Biomed Health Inform 2017; 21(4): 1095-1104. DOI: 10.1109/JBHI.2016.2594239.

[13] Tamiev D, Furman PE, Reuel NF. Automated classification of bacterial cell sub-populations with convolutional neural networks. PLoS ONE, 2020; 15: e0241200. DOI: 10.1371/journal.pone.0241200.

[14] Kang R, Park B, Eady M, Ouyang Q, Chen K. Single-cell classification of foodborne pathogens using hyperspectral microscope imaging coupled with deep learning frameworks. Sens Actuators 2020; 309: 127789. DOI: 10.1016/j.snb.2020.127789.

[15] Spahn C et al. DeepBacs: Bacterial image analysis using open-source deep learning approaches. bioRxiv Preprint. 2021. Source: <https://www.biorxiv.org/content/10.1101/2021.11.03.467152v1>. DOI: 10.1101/2021.11.03.467152.

[16] Smit JH, Li Y, Warszawik EM, Herrmann A, Cordes T. ColiCoords: A Python package for the analysis of bacterial fluorescence microscopy data. PLoS ONE 2019; 14 (6): 1-18. DOI: 10.1371/journal.pone.0217524.

[17] Stylianidou S, Brennan C, Nissen SB, Kuwada NJ, Wiggins PA. SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells. Molecular 2016; 102(4): 690-700. DOI: 10.1111/mmi.13486.

[18] Balomenos AD, Tsakanikas P, Aspridou Z, Tampakaki AP, Koutsoumanis KP, Manolakos ES. Image analysis driven single-cell analytics for systems microbiology. BMC Syst Biol 2017; 11(1). DOI: 10.1186/s12918-017-0399-z.

[19] Van Valen DA, Kudo T, Lane KM, Macklin DN, Quach NT, DeFelice MM, Maayan I, Tanouchi Y, Ashley EA, Covert MW. Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput Biol 2016; 12(11): e1005177. DOI: 10.1371/journal.pcbi.1005177.

[20] Smith PK, Kang AD, Kirby JE, Bourbeau P. Automated interpretation of blood culture gram stains by use of a deep convolutional neural network. J Clin Microbiol 2018; 56(3): e01521-17. DOI: 10.1128/jcm.01521-17.

[21] Koutsoumanis KP, Lianou A. Stochasticity in colonial growth dynamics of individual bacterial cells. Appl Environ Microbiol 2013; 79(7): 2294-2301. DOI: 10.1128/aem.03629-12.

[22] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 770-778. DOI: 10.1109/CVPR.2016.90.

[23] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In Book: Navab N, Hornegger J, Wells WM, Frangi AF, eds. Medical image computing and computer-assisted intervention – MICCAI 2015. Cham: Springer International Publishing Switzerland; 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.

[24] Weigert M. Content-aware image restoration: pushing the limits of fluorescence microscopy. Nature Methods 2018; 15(12): 1090-1097. DOI: 10.1038/s41592-018-0216-7.

[25] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. Proc IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 1125-1134. DOI: 10.1109/CVPR.2017.632.

[26] Schmidt U, Weigert M, Broaddus C, Myers G. Cell detection with star-convex polygons. In Book: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, eds. Medical image computing and computer assisted intervention – MICCAI 2018. Cham: Springer Nature Switzerland AG; 2018: 265-273. DOI: 10.1007/978-3-030-00934-2_30.

[27] Mandal S, Uhlmann V. Splinedist: Automated cell segmentation with spline curves. IEEE 18th Int Symposium on Biomedical Imaging (ISBI) 2021: 1082-1086. DOI: 10.1109/ISBI48211.2021.9433928.

[28] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 7263-7271. DOI: 10.1109/CVPR.2017.690.

[29] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997; 8: 1735-1780. DOI: 10.1162/neco.1997.9.8.1735.

[30] Gvozdev YA, Zimina TM, Kraeva LA, Hamdulaeva GN. Image recognition of juvenile colonies of pathogenic microorganisms in the culture based microbiological method implemented in bioMEMS device for express species identification. IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conf (EIConRusNW) 2016: 759-763. DOI: 10.1109/EIConRusNW.2016.7448292.

[31] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 2017; 39(6): 1137-1149. DOI: 10.1109/TPAMI.2016.2577031.

[32] Szegedy C, et al. Going deeper with convolutions. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2015: 1-9. DOI: 10.1109/CVPR.2015.7298594.

[33] Zhou K, Yang Y, Qiao Y, Xiang T. Domain Adaptive Ensemble Learning. IEEE Trans Image Process 2021; 30: 8008-8018. DOI: 10.1109/TIP.2021.3112012.

[34] Li H, Ng J, Natsev P. EnsembleNet: End-to-end optimization of multi-headed models. arXiv Preprint. 2019. Source: <https://arxiv.org/abs/1905.09979>. DOI: 10.48550/arXiv.1905.09979.

[35] Mikhalkova MA, Yachnaya VO, Malashin RO. Comparative analysis of convolutional neural networks and vision transformer on classification of images containing homogenous microstructures. Wave Electronics and Its Application in Information and Telecommunication Systems (WECONF) 2023: 1-6. DOI: 10.1109/WECONF57201.2023.10148032.

[36] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conf on Computer Vision and Pattern Recognition 2009: 248-255 DOI: 10.1109/CVPR.2009.5206848.

[37] Kingma DP, Ba J. Adam: A method for stochastic optimization. 3rd Int Conf for Learning Representations 2015.

[38] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. MobileNetV2: Inverted residuals and linear bottlenecks. IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 4510-4520. DOI: 10.1109/CVPR.2018.00474.

[39] Mingxing T, Le Quoc V. EfficientNet: Rethinking model scaling for convolutional neural networks. Int Conf on Machine Learning 2019: 6105-6114.

[40] Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014; 15(56): 1929-1958.

[41] Zhu Z, Ren Z, Wang SH, Górriz JM, Zhang YD. RDNet: ResNet-18 with dropout for blood cell classification. In Book: Vicente JMF, Álvarez-Sánchez JR, de la Paz López F, Adeli H, eds. Artificial intelligence in neuroscience: Affective analysis and health applications. Cham: Springer Nature Switzerland AG; 2022: 136-144. DOI: 10.1007/978-3-031-06242-1_14.

[42] Hua SB, Lu AX, Moses AM. CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning. NeurIPS 2021 Learning Meaningful Representations for Life (LMRL) Workshop 2021: 1-7. DOI: 10.48550/arXiv.2111.11646.

## *Appendix A*

### Choosing a classifier architecture

Convolutional neural networks show advantage in recognizing images with homogeneous microstructure compared to transformers [35].

Further analysis of related works does not give unambiguous recommendations on architecture choice, as they did not consider the problem of classification of a large number of bacterial species that were not prepared as microscopic slides (preparations). Hence, four widely used convolutional neural networks are further considered to assess their applicability in bacteria classification: MobileNetV2 [35], ResNet-18 and ResNet-50 [22], EfficientNet-B7 [36]. A hybrid ResNet-SVM classifier was also tested, as SVM is successfully applied in related works, such as in [7, 9], and we expected it to provide good generalization abilities on our comparatively small dataset.

The advantages of MobileNetV2 are its "compactness" in terms of the number of weights; it also has higher computational speed compared to the other architectures considered [35]. Resnet-18 is also relatively compact network, while Resnet-50 is wider and deeper. The EfficientNet-B7 is the largest considered architecture that belongs to the class of EfficientNet architectures, which reduce computational complexity by scaling all dimensions of depth, width and resolution [36].

The support vector machine was trained on pre-extracted image features extracted from the last convolutional layer of the 4th Res-block of the ResNet-18 after applying the global pooling. The ResNet-18's weights reaching the best accuracy of 30.3 % on the test set were chosen. However, ResNet-SVM classification accuracy dropped to 26.4 %. If we used ImageNet pretrained features, then SVM accuracy dropped even further.

The resolution of the input images was adjusted as a hyperparameter in accordance with the available computing resources, and for all models was set as 512×512 pixels. The classification results are shown in Tab. 4.

*Tab. 4. Classification accuracy of different architectures (averaging over three training runs), input image size is 512×512 pixels*

| Model | Acc. (val), % | Acc. (test), % | MMA (test), % |
|---|---|---|---|
| MobileNetV2 | 97.5 | 27.9 | 33.02 |
| ResNet-18 | 97.2 | 30.15 | **36.9** |
| ResNet-50 | 98.7 | 29.9 | 31.4 |
| EfficientNet-B7 | 95.5 | **31.6** | 32.28 |
| Resnet-SVM | - | 26.4% | - |

It should be reminded, that the validation set consists of examples of the same strains that are in the training set, while the test set contains only new strains. As can be seen from Tab. 4, all the considered models achieve a rather high classification accuracy at the training and validation phases (over 95 %). However, the accuracy of these models is much lower during the testing phase. The high classification accuracy on the validation set proves that examples within the same strain have lower visual variability, which allows algorithms to extract features and dependencies in them. At the same time, low accuracy on the test set indicates the significant visual variability among strains of the same species (Fig. 1).

According to results obtained, the EfficientNet-B7 reaches the highest classification accuracy. However, the computing resources available allow us to increase the size of the input image for other architectures, which, as will be shown below, improves recognition accuracy. Such image scaling was not possible for EfficientNet-B7, so all further experiments are carried out with the ResNet-18 model, which achieved best MMA and second best average test accuracy.

### Color mode and image resolution effect on classification accuracy

As can be seen from Fig. 1, images of different bacterial species and strains have different colors; accordingly, this feature must be excluded by converting all images in the dataset to grayscale. In addition, there are also differences in saturation, brightness and contrast of the images. To prevent the classifier overfitting on these parameters we used brightness and contrast augmentation.

To estimate color mode effect on the classification accuracy, the ResNet-18 model was trained and tested on the original (colored) dataset and on a modified (grayscale) one. The results obtained on the grayscale dataset applying

brightness and contrast augmentation exceed the ones on the colored dataset by 10 %: average accuracy on the test set is 34.47 % and MMA on the tests set reaches 36.9 %.

Optimal resolution choice of the input image may help to increase recognition accuracy. For this purpose, we trained three models of the ResNet-18 processing input images of various sizes: 512×512, 704×704, 1024×1024 pixels. Note, that an image of any resolution can be fed to the network since the ResNet uses a global average pooling operation before the fully connected layer, which works with a tensor of any size. The higher resolution provides better accuracy, e.g., while the MMA on the 512×512 pixel images is 30.3 %, on the 1024×1024 pixel images it reaches 36.9 %. Thus, for further experiments, the resolution of 1024×1024 pixels for the input images is used.

*Authors' information*

**Valeria Olegovna Yachnaya,** (b. 1997) graduated from the Saint Petersburg State University of Aerospace Instrumentation in 2021, majoring in Computer Science and Engineering. Currently she is a graduate student at the Saint Petersburg State University of Aerospace Instrumentation and works as the assistant research worker at the Pavlov Institute of Physiology of RAS. Research interest is computer vision. E-mail: *tamimio.yvo@hotmail.com*

**Maria Anatolievna Mikhalkova,** (b. 1997) graduated from the Saint Petersburg State University of Aerospace Instrumentation in 2021, majoring in Computer Science and Engineering. Currently she works as the assistant research worker at the Pavlov Institute of Physiology of RAS. Research interest is computer vision.
E-mail: *mikhalkova_maria@infran.ru*

**Roman Olegovich Malashin,** (b. 1987) graduated from the Saint Petersburg State University of Aerospace Instrumentation in 2011. He received a PhD degree at University of Information Technology, Mechanics and Optics in 2014. The title of the PhD thesis was "Structural analysis of images of 3-D scenes". From 2017 he is the head of the Artificial Intelligence and Neural Networks Group at the Pavlov Institute of Physiology of RAS. Research interest lies in the fields of computer vision, machine learning and artificial intelligence. E-mail: *malashinroman@mail.ru*

**Vadim Rostislavovich Lutsiv,** (b. 1954) graduated from the Leningrad Institute of Aerospace Instrumentation in 1977, majoring in Electronic Computers. He received a Doctor of Technical Science degree by the decision of the dissertation council at the Saint Petersburg State University of Aerospace Instrumentation in 2012. E-mail: *vluciv@mail.ru*

**Lyudmila Aleksandrovna Kraeva,** (b. 1963) graduated from Dnepropetrovsk Medical Institute in 1986 with a degree in Microbiology, Epidemiology, Hygiene. Her doctoral dissertation in Microbiology (medical sciences) on the topic "Biological foundations of the development of new technologies for the diagnosis and monitoring of diphtheria" was defended in 2011. Since 2015 she is the head of the Laboratory of Medical Bacteriology at the St. Petersburg Pasteur Institute. E-mail: *lykraeva@yandex.ru*

**Galina Nikolaevna Khamdulayeva,** (b. 1967) graduated from Dagestan State University in 1991 with a degree in Biology. Presently, she working as a junior researcher in the Laboratory of Medical Bacteriology at the St. Petersburg Pasteur Institute. She is doing her PhD thesis on the topic "Development of a method for express identification of infectious agents based on microtechnologies". E-mail: *Galina.khamdulaeva@yandex.ru*

**Vitaliy Evgenievich Nazarov,** (b 1964) graduated from Military Medical Academy named after S.M. Kirov in 1987. He received a Doctor of Medical Science degree by the decision of the dissertation council at the Military Medical Academy named after S.M. Kirov in 2002. Currently he works as the Professor of Surgery department named after I.I. Grekov of North-Western State Medical University named after I.I. Mechnikov.
E-mail: *VENazarov@yandex.ru*

**Vladimir Petrovich Chelibanov,** (b. 1952) graduated from Tambov State Technical University in 1975. In the period from 1976 to 1980, he studied at the graduate school of the Leningrad Technological Institute, specializing in Chemical Physics. Upon completion of graduate school, he defended his thesis on the topic: "Study of physico-chemical processes during the photolysis of halogen-substituted saturated hydrocarbons." Since 1993, he has been the head of the Russian instrument-making enterprise OPTEC JSC. His scientific interest lies in the development of composite optically active materials, nonlinear optics and optical signal processing technologies.
E-mail: *chelibanov@gmail.com*