Journal of Siberian Federal University. Biology 2025 18(1): 132–147

~ ~ ~

EDN: GETAZO

УДК 004.93.032.26’1:519.688:576.89

A Neural Network Algorithm for Identifying Monogeneans of the Order Dactylogyridea

Alexey V. Kazarnikov[a*], Yuliya V. Stepanova[b] and Anna V. Kazarnikova[b]

[a]Southern Mathematical Institute – branch of Vladikavkaz Scientific Centre of the RAS Vladikavkaz, Russian Federation

[b]Southern Scientific Centre of RAS Rostov-on-Don, Russian Federation

Received 19.09.2023, received in revised form 19.03.2025, accepted 20.03.2025

Abstract. Aquaculture as one of the fastest growing food-producing sectors has given rise to numerous fish farms which often lack the laboratory facilities and qualified professionals to accurately diagnose disease and provide appropriate treatments. Monogeneans are parasitic worms that belong to the phylum Platyhelminthes. Some of them can cause mass mortality of fish in both natural water bodies and aquaculture settings. This paper presents a neural network algorithm designed to accurately identify monogeneans of the order Dactylogyridea using digital photographs taken through the ocular lens of the light microscope with a smartphone camera. The lowered equipment requirements make this method suitable for use on fish farms. We frame the identification of Dactylogyridea as a binary classification problem and train the VGG-16 convolutional neural network to classify images of these monogeneans. To enhance our dataset, we apply data augmentation techniques that artificially increase the number of training examples and simulate variations in microscope illumination levels, image under- / overexposure or parasite discolouration on a slide. Our recognition algorithm achieves a classification accuracy of 98.8 % for the elements of both testing and validation sets. The results obtained in this study are of practical value since many species of the order Dactylogyridea can cause lethal diseases in fish. This method can improve diagnostics, treatment and disease prevention in aquaculture. Additionally, its simplicity is particularly advantageous for novice specialists.

Keywords: monogeneans, Dactylogyridea, fish parasites identification, convolutional neural network.

* Corresponding author E-mail address: kazarnikov@gmail.com

ORCID: 0000-0002-5956-2791 (Kazarnikov A.); 0000-0003-2903-0361 (Stepanova Y.); 0000-0002-3110-3120 (Kazarnikova A.)

Acknowledgements. The biological material for this work was obtained from research conducted by the Southern Scientific Centre of the Russian Academy of Sciences (SSC RAS) in 2019–2022 in the Don River Delta, Svinoe Girlo Strait, and the eastern part of the Taganrog Bay. The research carried out by Anna V. Kazarnikova and Yulia V. Stepanova was financially supported by the Ministry of Education and Science of the Russian Federation (agreement No. 075–15–2024–528 dated April 24, 2024, for the implementation of major scientific projects (MSP) in priority areas of scientific and technological development). Additionally, we would like to thank Cand. Sc. (Biology) E. V. Shestakovskaya for advice and constructive critique in the course of study as well as the three anonymous reviewers for their valuable comments which helped us to significantly improve the quality of the manuscript.

Citation: Kazarnikov A. V., Stepanova Y. V., Kazarnikova A. V. A neural network algorithm for identifying monogeneans of the order Dactylogyridea. J. Sib. Fed. Univ. Biol., 2025, 18(1), 132–147. EDN: GETAZO

Разработка нейросетевого алгоритма для определения принадлежности моногеней к отряду Dactylogyridea

А. В. Казарников[а], Ю. В. Степанова[б], А. В. Казарникова[б]
[а]Южный математический институт – филиал Владикавказского научного центра РАН Российская Федерация, Владикавказ
[б]Южный научный центр Российской академии наук Российская Федерация, Ростов-на-Дону

Аннотация. Развитие аквакультуры ведёт к появлению новых рыбоводных хозяйств, на которых зачастую отсутствуют рыбохозяйственные лаборатории и специалисты, способные поставить диагноз о заболевании и выбрать соответствующие методы его лечения. Моногенеи относятся к типу плоских червей (Platyhelminthes). Некоторые из них способны приводить к массовой гибели рыб в естественных водоёмах и при выращивании в аквакультуре. В работе представлен нейросетевой алгоритм, позволяющий с высокой точностью идентифицировать представителей отряда Dactylogyridea по фотографиям, сделанным через окуляр светового микроскопа камерой обычного смартфона. Такие требования к оборудованию делают возможным применение данного подхода на рыбохозяйственных предприятиях. Проблема распознавания представителей отряда Dactylogyridea сводится к задаче бинарной классификации, и для распознавания изображений моногеней проводится обучение свёрточной нейронной сети VGG-16. Для искусственного увеличения количества обучающих примеров, имитации эффектов засветки и затемнения изображений, моделирования различных уровней освещённости исследуемых под микроскопом паразитов использованы методы аугментации данных. Точность классификации составила 98,8 % на элементах тестового и валидационного множеств. Полученные результаты имеют практическую ценность, так как среди представителей рассматриваемого отряда много возбудителей опасных болезней, приводящих к гибели рыб. Разработанный подход может упростить работу начинающих специалистов, а также способствовать более быстрой диагностике заболеваний и разработке методов их профилактики и лечения.

Ключевые слова: моногенеи, дактилогириды, определение паразитов рыб, свёрточная нейронная сеть.

Благодарности. Биологический материал, используемый в данной работе, был получен в результате исследований, проводимых ЮНЦ РАН в 2019–2022 годах в дельте реки Дон, пр. Свиное гирло, и восточной части Таганрогского залива. Работа А. В. Казарниковой и Ю. В. Степановой была выполнена при финансовой поддержке Минобрнауки России (соглашение № 075–15–2024–528 от 24.04.2024 г. на реализацию КНП по приоритетным направлениям научно-технологического развития). Кроме того, мы хотели бы поблагодарить канд. биол. наук Е. В. Шестаковскую за полезные советы и конструктивное обсуждение во время проведения данного исследования, а также трёх анонимных рецензентов за их ценные замечания, которые помогли нам значительно улучшить качество данной работы.

Цитирование: Казарников А. В. Разработка нейросетевого алгоритма для определения принадлежности моногеней к отряду Dactylogyridea / А. В. Казарников, Ю. В. Степанова, А. В. Казарникова // Журн. Сиб. федер. ун-та. Биология, 2025. 18(1). С. 132–147. EDN: GETAZO

Introduction

Aquaculture is the fastest-growing sector of agriculture (Food and Agriculture Organization of the United Nations [FAO], 2022). According to the Food and Agriculture Organisation of the United Nations (FAO, 2022), global aquaculture production, including fish, crustaceans, molluscs and other aquatic animals (excluding marine mammals and reptiles), reached 88 million tonnes in 2020, which is 49 % of the total fishery production (178 million tonnes). Over 157 million tonnes (89 %) were used for human consumption.

Different forms of ownership (private or governmental) and various fish farming methods nowadays aggravate the ecological and epizootic conditions (Butko et al., 2017). Poor management of production processes on fish farms and water pollution cause outbreaks of disease of various aetiology in aquaculture. Disease is commonly recognised as a serious threat to the commercial success of aquaculture. Direct loss of fish may be up to 100 % (Feist et al., 2019).

Aquaculture has given rise to numerous fish farms which often do not have laboratories and professionals well qualified to diagnose disease and provide appropriate treatments. These professionals need education and skills to work with the light microscope. Thus the fish farmer is burdened with another responsibility in addition to regular activities (breeding, rearing and harvesting fish).

Monogeneans (Monogenea (Van Beneden, 1858) Bychowsky, 1937) are parasitic worms, which belong to the phylum Platyhelminthes. Their systematic position is shown in Fig. 1. Their definitive hosts are fish. Most monogeneans parasitise gills, skin, fins, nasal and oral cavities (Gaevskaya, 2004). They feed on host epithelial cells, mucous secretions from skin, and blood. Some monogeneans (for example, members of the genera *Gyrodactylus* and *Dactylogyrus*) can cause mass mortality of fish in natural water bodies

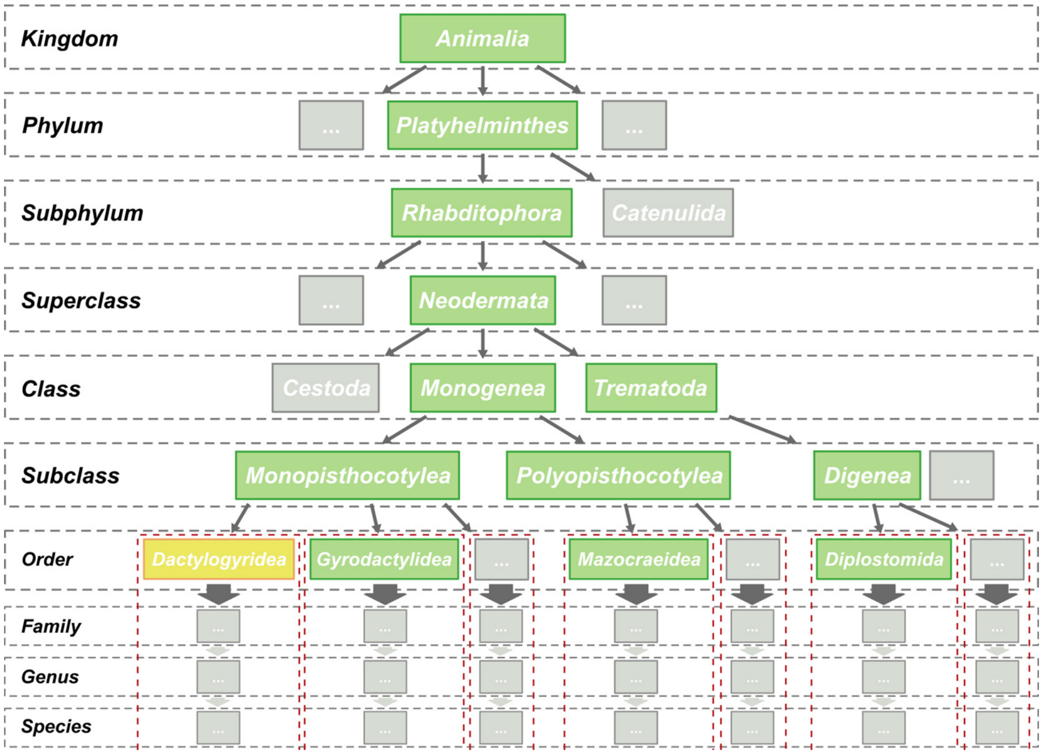

Fig. 1. The systematic position of the class Monogenea and the order Dactylogyridea (according to the taxonomic database World Register of Marine Species (WoRMS, 2023). Biological material was collected for the orders in green boxes. The boundaries between the orders are given in red. These boundaries illustrate the accuracy level of the recognition algorithm developed in the paper. The smaller the subtree size inside the border, the higher the recognition accuracy

and aquaculture (Bauer et al., 1981). In this case, juveniles are affected most severely. Adult fish are less susceptible to the disease and act as parasite carriers.

Basic techniques for collecting and processing the biological material of monogeneans were developed long ago (Gusev, 1983). A small amount of mucus is scrapped off gills, fins or body surface of fish and examined under a microscope with magnification allowing for a clear view of individual parasites. Most monogenean species can be identified using morphological criteria, mainly the shape and size of the elements of the attachment disc (Pugachev et al., 2009; Vignon, 2011a; Strona et al., 2014; Yousef Kalafi et al., 2016). However, morphological features are highly variable both within and between species. Therefore, identification of parasite species via microscopic examination requires special training and ample experience.

The identified parasite is treated using a method generally recommended for combating monogeneans. The choice of the method depends on the country's regulations and drugs available. In each particular case it also depends on the age of fish and the cultivation method.

The aim of this study is to develop a neural network algorithm for recognising Dactylogyridea in digital photographs. A distinctive feature of the present research is that we impose minimal requirements on the quality of the image and assume that it was obtained by photographing through the ocular lens of the light microscope with a smartphone camera. The reduced equipment requirements make our method suitable for use on fish farms where expensive apparatus is not available. The images obtained this way do not contain enough information to identify monogenean species by morphological characteristics. However, it is still possible to classify parasites at a higher taxonomic level (between orders).

Materials and methods

Data collection

For an individual parasite analysis, 578 cyprinids (carp, bream, ram, goldfish) and sanders (the family Percidae) fish were taken. The study for infection with monogenetic flukes was carried out using standard methods (Bykhovskaya-Pavlovskaya, 1985). Gill covers were removed and all gill arches were excised with scissors. Arches were examined under a binocular stereoscopic microscope MBS 10 at 8 × 4 magnification. A sample of mucus was scraped off each arch with a scalpel. A drop of water was added to the mucus to reduce viscosity and the substance was placed on a glass slide and examined under a microscope for presence of monogeneans. Taxonomic affiliation was determined using "Key to the parasites of freshwater fish of the USSR" (Bauer, 1985). Taxa nomenclature was given according to the World Register of Marine Species (WoRMS, 2023). The Mikmed-5 microscope (4 × 0.12 × 10) was used to analyse the objects in transmitted light by the light field method. Digital photographs of parasites were obtained using Nokia 7.2 and Samsung A50 smartphones.

Architecture of artificial neural network

To classify the images of monogeneans, the VGG-16 convolutional neural network (an artificial neural network with several convolutional layers) was used. It was proposed by Simonyan and Zisserman (2015) as an improvement to the AlexNet architecture (Krizhevsky et al., 2017), which earlier showed good results in image classification at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2010[1]. The main improvement in the VGG-16 architecture as compared to AlexNet was a larger number of layers and a lower resolution of the convolution window. This improvement made it possible to optimise the number of weights and at the same time achieve a higher accuracy of image classification compared to AlexNet at the ILSVRC 2014 competition (Nikolenko et al., 2018).

As an input, the VGG-16 network accepts an RGB image with a fixed resolution of 224 × 224 pixels. The network consists of five convolutional modules followed by three fully connected neuron layers (see Fig. 2). Each convolutional module contains two or three convolutional layers with 3 × 3 windows, followed by a subsampling (max pooling) layer with a 2 × 2 window. Thus, after passing through each module, the resolution of the feature map is halved. The first two fully connected layers consist of 4096 neurons each, and the number of neurons in the output layer of the network corresponds to the number of input data classes. In total, the network contains 16 neuron layers. Nonlinear activations are defined by the ReLU function in all internal layers, and by the SoftMax function in the output layer. Precise mathematical definitions of the above-mentioned activation functions, as well as convolutional, subsampling and fully-connected layers can be found in (Goodfellow et al., 2016) or (Nikolenko et al., 2018). The output values of the network $0 \leq p_j \leq 1$, $p_1 + \ldots + p_N = 1$, are interpreted as the model's degree of confidence that the input image belongs to each of N predefined data classes. In this work, the number of output classes is equal

[1] In this competition, various image recognition models compete annually in classification of objects from the ImageNet database, which contains several millions of annotated images (Russakovsky et al., 2015).

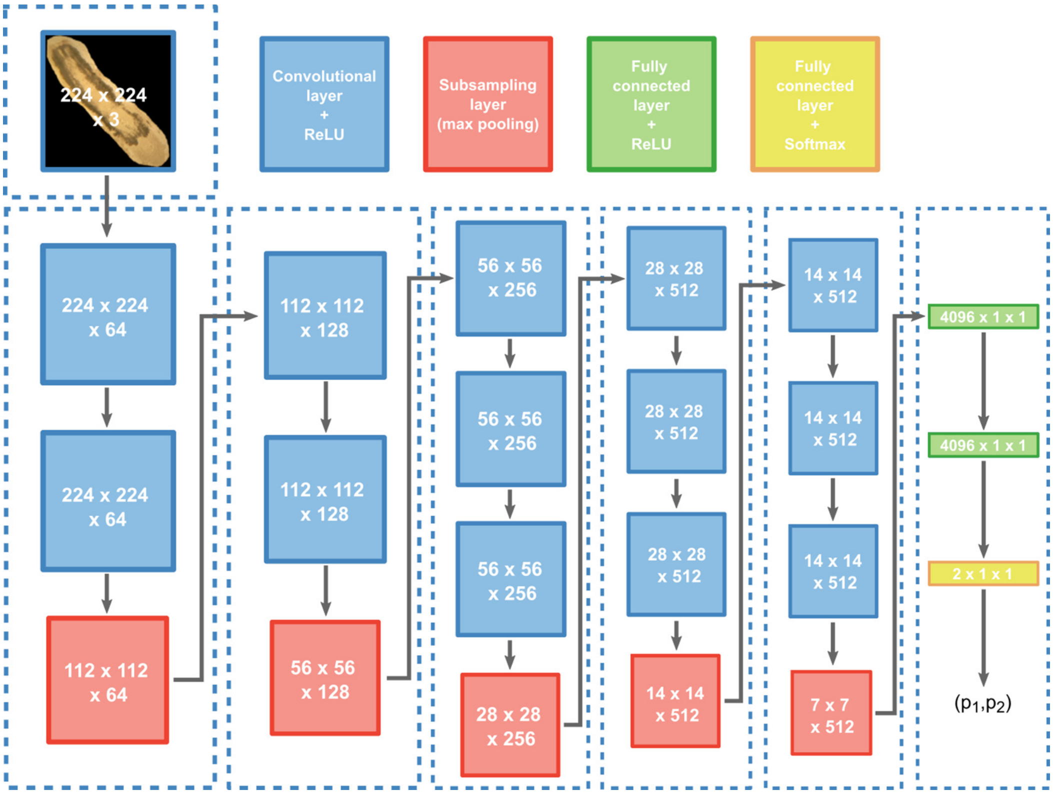

Fig. 2. Architecture of VGG-16 convolutional neural network. It consists of five convolutional modules and three fully connected neuron layers. As input, the network receives an image of the parasite after the data augmentation procedure. The output is a pair of values (p_1, p_2), $p_1 + p_2 = 1$, which express the model's degree of confidence that the input data belong to two given classes

to two (class 0: the parasite belongs to the order Dactylogyridea, class 1: it does not belong to the order Dactylogyridea).

Data pre-processing and augmentation

We initiate the image pre-processing by manually removing the background from the images. Subsequently, the training data for the neural network is divided into two classes. The first class comprises photographs of members of the order Dactylogyridea, while the second class includes examples of "other" parasites that do not belong to this order. Consequently, the task of recognising members of the order Dactylogyridea is framed as a binary classification problem (Viola & Jones, 2001; Isukapalli et al., 2006; Siradjuddin et al., 2020).

The total amount of collected biological material is given in Table 1.

In collecting data for the first class, it was essential to capture the full extent of variability within the order Dactylogyridea. We considered the following factors (see also Fig. 3):

1. **Body contraction.** Monogeneans are motile parasites that can quickly change body shape. While moving, their elastic bodies can be considerably deformed. Therefore, a training data set should contain a variety of examples of different contraction degree of worm bodies.

2. **Species variability.** Dactylogyridea are relatively small worms with 2 pairs of eyes at the anterior end of the body. At the posterior end, there is an attachment disk with 14 marginal hooks, 2–4 median hooks and 1–3 connecting plates

Table 1. Biological material used to train the neural network model

Members of the order Dactylogyridea	Number of samples
Genus *Ancyrocephalus* (family Ancyrocephalidae)	36
Genus *Dactylogyrus* (family Dactylogyridae)	94
Genus *Ligophorus* (family Ancyrocephalidae)	42
TOTAL	172
Members of other orders	
Order Diplostomida (class Trematoda)	9
Order Gyrodactylidea (class Monogenea)	36
Order Mazocraeidea (class Monogenea)	10
TOTAL	55

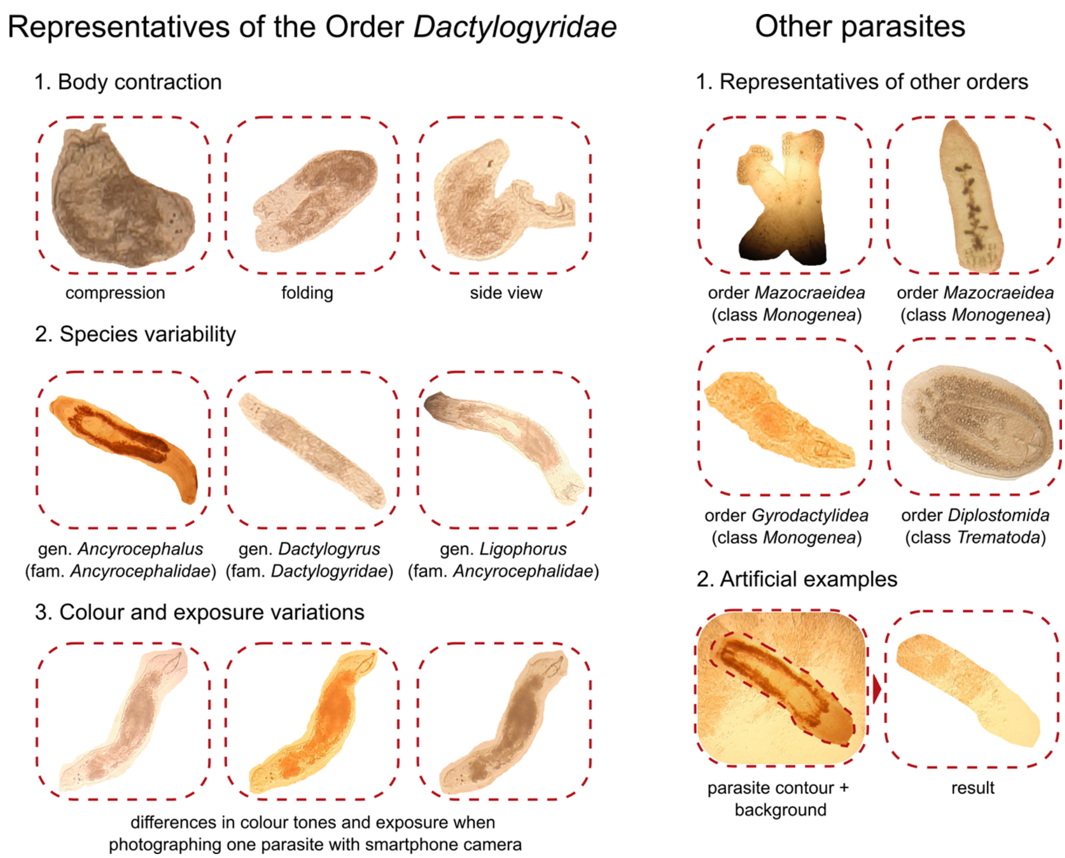

Fig. 3. Organisation of a training set for the neural network. The first data class includes examples of members of the order Dactylogyridea. The second class includes images of parasites that do not belong to this order, as well as artificially created examples

which may be secondarily absent (Bauer, 1985). The covers frequently form an annular folding and have cuticular spines for better attachment to the host. Despite the fact that all members of the order have common morphology, specific features (the shape of the attachment disk, the number or size of the hooks, the type of pigmented eyes, and others) differ between different families of monogeneans within to the order. The shape and structure of the attachment hooks and the copulatory organ serve as distinctive features for species identification. The training set included families most common in freshwater bodies in Russia: Dactylogyridae (Bychowsky, 1933) and Ancyrocephalidae (Bychowsky, 1937).

3. **Colour and exposure variations.** When photographing a parasite through the ocular lens of a microscope, the automatic adjustment of the camera parameters by the algorithms of the mobile operating system can produce images of different colour and tone, as well as under- or overexposed images. In addition, ingress of blood or dirt into the mucus scrape may lead to discolouration of the parasite. To reduce the effect of these factors on accuracy of the algorithm, we performed data augmentation.

The second class of training data contains photographs of monogeneans (orders Gyrodactylidea, Mazocraeidea) and trematodes (order Diplostomida). Members of these orders are shown in Fig. 3. The training set included a sufficient number of members of the order Gyrodactylidea as their morphology closely resembles that of the order Dactylogyridea. To achieve balance between the data classes, artificial examples were generated to ensure that the number of examples in each class was comparable. To create these artificial examples, randomly selected regions corresponding to the contours of the order Dactylogyridea parasites were cut out of a rectangular fragment of the light field.

Since the collection of biological material is a long and laborious process, lack of training data is a common problem in research related to machine learning methods in aquaculture (see Table 2). To cope with it, it is necessary to perform data augmentation to increase artificially the number of training examples before teaching algorithms to a neural network (Shorten & Khoshgoftaar, 2019). We also used this procedure to simulate possible variations of microscope illumination levels and image under- / overexposure when photographing a parasite through the ocular lens of the microscope, as well as parasite discoloration on a slide. The proposed data augmentation scheme is shown in Fig. 4 and consists of the following steps:

1. Colour channel distortion: random perturbation of each channel by uniform random noise (with probability $p^{1}_{\mathrm{colour}} = 0.4$), or channel distortion along the principal intensity components of RGB channels (with probability $p^{2}_{\mathrm{colour}} = 0.4$), or no change (with probability $p^{3}_{\mathrm{colour}} = 0.2$).

2. Horizontal flipping (with probability $p_{\mathrm{flip}} = 0.5$).

3. Rotation by a randomly chosen angle $\alpha_{\mathrm{rot}} \in [0,360)$.

Table 2. Size of training sets in studies on the application of machine learning algorithms in aquaculture

Paper	Studied problem	Number of samples
(Ahmed et al., 2022)	fungal fish diseases	266
(Hasan et al., 2022)	ichthyophthyriosis, red spot disease	90
(Yousef Kalafi et al., 2016)	monogenean classification (4 species)	80
(Ali et al., 2012)	distinguishing three species of *Gyrodactylus*	66
(Vignon, 2011b)	distinguishing three species of fam. Ancyrocephalidae	160
(Ali et al., 2011)	distinguishing nine species of *Gyrodactylus*	557

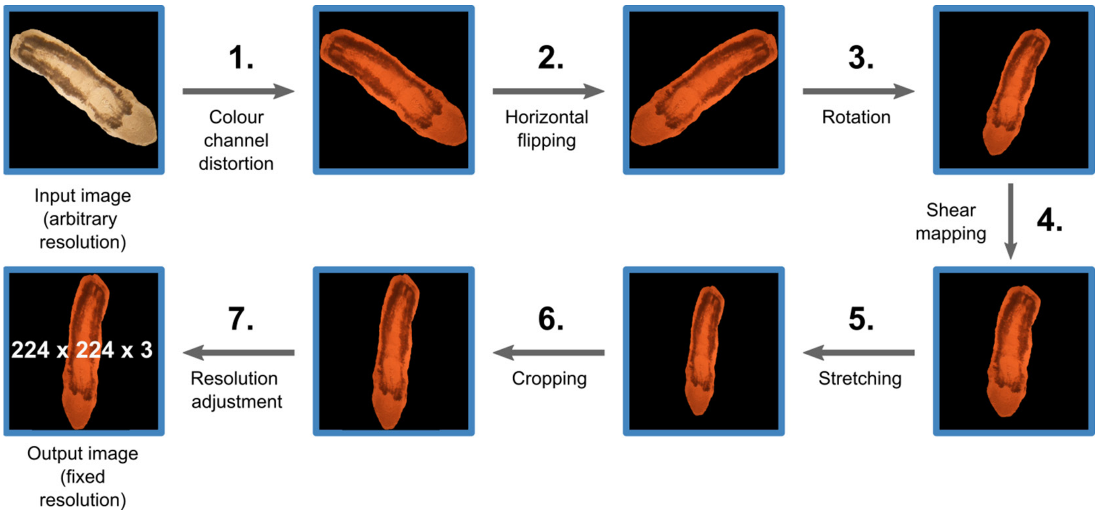

Fig. 4. Data augmentation procedure. The input is a digital image of the parasite at an arbitrary resolution, with the background manually removed. The output is a square image of 224 × 224 pixels

4. Shear mapping: $(x, y) \longrightarrow (x + \alpha^1_{shear} y, y + \alpha^2_{shear} x)$ with randomly chosen parameters α^1_{shear}, $\alpha^2_{shear} \in (0, 0.17)$.

5. Horizontal and vertical stretching with randomly chosen scaling factors α^1_{shear}, $\alpha^2_{shear} \in (0.87, 1.27)$.

6. Cropping (removing the frame from empty pixels) and bringing the image to a square shape.

7. Adjusting image resolution to 224 × 224 pixels.

Neural network training

When training a neural network model, no additional data pre-processing was performed, except for the manual removal of the background and augmentation procedure described in the previous section. We experimented with data centring (subtracting the global average of the intensity of RGB channels) and applying fixed-size image parts (Krizhevsky et al., 2017), but no positive effect of these procedures on the model performance or learning rate was registered. Regularisation of the neural network was implemented in the convolutional layers, with batch normalisation applied before the non-linearity (Ioffe & Szegedy, 2015). In the first two fully connected layers, dropout was used (Srivastava et al., 2014) with a probability $p_{drop} = 0.6$.

The training data (see Table 1) was divided into training, validation and testing sets in a ratio of 60:20:20, corresponding to 137, 45 and 45 images, respectively. For each element of the validation set, 100 augmented examples were generated, resulting in a total of 4,500 images. These augmented examples were used to assess the model's generalisation capability during the training process. To optimise performance, data augmentation on the validation set was performed only once, prior to the start of training.

The VGG-16 convolutional neural network was trained using the TensorFlow 2.9.3 library. The program code was implemented in Python 3.9, with the OpenCV library tools employed for image processing. Data augmentation was conducted in asynchronous mode on the CPU, while the model itself was trained on an Nvidia RTX A5500 graphics card. The data were divided into fixed-size batches of augmented examples, each batch consisting of 750 images of members of the order Dactylogyridea, 450 images of members of other orders, and 300 artificial examples, making a total of 1,500 images per batch. During the data augmentation, the images were randomly sampled from the respective classes of the training dataset.

The first batch of augmented data was created before training began. Once this batch was ready, the training process was immediately started on the GPU. Simultaneously, a CPU thread was initiated to asynchronously prepare the next batch of augmented data. Upon the termination of the thread, the data were changed, allowing the training to continue with the newly prepared batch. The thread was then restarted to prepare the subsequent batch of data. This process ensured a continuous and efficient training workflow by running in parallel data preparation and model training. The complete code for implementation is available at https://github.com/AlexeyKazarnikov/NNParasiteRecognitionCode.

The model was trained using the stochastic gradient descent method with the learning rate $\eta = 10^{-3}$, the L_2-regularisation parameter $\lambda = 5 \times 10^{-4}$, and the Nesterov gradient weight parameter $\gamma = 0.9$ (Simonyan & Zisserman, 2015). Cross-entropy was used as the loss function. In the course of training, we reduced the learning rate by a factor of ten whenever the model's performance on the validation set did not improve over several epochs. We finished the training process when the model achieved the accuracy of 99.9 % on the training set and approximately 92 % on the validation set, with no further improvement observed over 50 epochs following the last reduction of the learning rate.

Results

The neural network model testing on the elements of the validation and testing sets (90 examples in total) was carried out. For each image, 100 augmented examples were created and model predictions for each of them were calculated. A conclusion whether the parasite belongs to the order Dactylogyridea was based on the averaged value of predictions. Classification accuracy was 98.8 percent. It means that only one image was incorrectly classified, when the pigment eyes, a distinctive feature of the order Dactylogyridea, were difficult to distinguish due to the poor quality of the photograph. The confusion matrix, which visualises the performance of this recognition algorithm on the validation and training sets, is shown in Table 3. Fig. 5 shows the evolution of the accuracy and loss functions during network training.

Table 3. Confusion matrix visualising the performance of the recognition algorithm on the elements of the validation and testing sets

		Predicted	
	Total 90	Member of the order *D.* 67	Not a member of the order *D.* 23
Actual	Member of the order *D.* 68	True positive 67	False negative 1
Actual	Not a member of the order *D.* 22	False positive 0	True negative 22
	Accuracy 98.8 %		

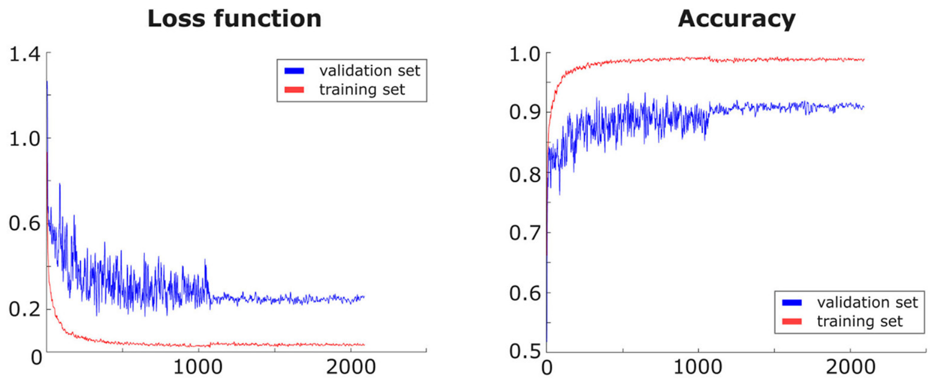

Fig. 5. Evolution of the accuracy and loss functions on the training and validation sets during the training process. Function values were smoothed before plotting using moving averages over a sliding window of 5 epochs

Discussion

Machine learning has been successfully employed in aquaculture for an extended period. Review papers (Awalludin et al., 2020; Yang et al., 2021) summarise current advances in the automation of routine tasks within the field, while other studies (Li et al., 2022; Pauzi et al., 2021) focus specifically on the application of machine learning in diagnosing fish diseases.

Machine learning algorithms can be integrated into automated systems capable of detecting outbreaks of fish disease in pools or aquaria. Such systems can alert users to problems and in some cases automatically administer necessary treatments or preventive measures. Li et al. (2023) proposed a system for detecting several harmful parasites in aquarium goldfish (*Carassius auratus*), including flatworms (*Gyrodactylus kobayashii*: Monogenea), aquatic crustaceans (*Argulus japonicus*: Ichthyostraca) and ciliates (*Ichthyophthirius multifiliis*: Oligohymenophorea). Data captured by a digital camera are processed in real-time by a convolutional neural network that detects parasites on the bodies of passing fish. When a monitoring system identifies that pest levels have surpassed a predetermined threshold, it can automatically dispense medicine or treatments through an automatic feed system to control the infestation. Waleed et al. (2019) developed a method for automatic diagnosis of three different fish diseases: epizootic ulcerative syndrome (red spot disease, caused by microscopic fungi), ichthyophthiriosis (caused by the ciliate *I. multifiliis*) and columnariosis (caused by the bacterium *Flavobacterium columnaris*: Flavobacteria). It consists in analysing images of passing fish by a computerised system that takes into account data from sensors measuring the water parameters in the pool. This information is sent to a computer, where it is analysed by a neural network algorithm. When a disease outbreak is detected, an alert is sent to the user's mobile phone.

Although automated systems are able to effectively detect fish diseases, they require expensive equipment and maintenance. Thus, laboratory diagnostics is another important application area for machine learning and image processing methods in aquaculture. Normally, digital photographs of the body surface and microscope images are used as input data for the algorithms. Ahmed et al. (2022) trained a linear classifier model (support vector machine) to detect symptoms of fungal infections in photographs of the body surface of salmon. Hasan et al. (2022) applied convolutional neural networks to detect fish infected with ichthyophthiriosis and red spot disease. Park et al. (2007) introduced a statistical algorithm for using microscope images to detect three types of disease of the Asian paralicht (*Paralichthys olivaceus*): ichthyophthiriosis, trichodiniosis (caused by peritrichous ciliates of the family Trichodinidae: Oligohymenophorea) and scuticociliatosis (caused by parasitic ciliates of the subclass Scuticociliatia: Oligohymenophorea). Zhan et al. (2020) employed image segmentation methods to find suitable morphological features for distinguishing between four species from the genus *Myxobolus*: Myxosporea.

When performing species identification of monogeneans, trained machine learning algorithms can be used to classify species using a set of morphometric characteristics of the attachment organ. Vignon (2011b) statistically analysed applicability of various characteristics of the parasite attachment apparatus for species identification, and carried out a classification of four monogenean species from the family Ancyrocephalidae. Yousef Kalafi et al. (2016) classified four monogenean species (*Sinodiplectanotrema malayanum*, *Trianchoratus pahangensis*, *Metahaliotrema mizellei* and *Metahaliotrema* sp.) using the automated identification technique based on the K-nearest neighbours method (KNN). This work used various geometric characteristics of attachment hooks as input data. Ali et al. (2012) used several machine learning classifiers, including a fully connected neural network, to distinguish between three *Gyrodactylus* species (*G. derjavinoides*, *G. salaris*, and *G. truttae*) that parasitise salmon. The authors applied Active Shape Models to the segmented areas of the image which corresponded to attachment hooks and used the Principal Component Analysis (PCA) to reduce the dimensions of the resulting sets of points. The output data were used as the input to the machine learning classifier. Ali et al. (2011) analysed the applicability of different morphometric features of attachment hooks for distinguishing between nine parasite species from the genus *Gyrodactylus*, including the pathogenic species *G. salaris*, and compared performance of several linear classifier models applied to this task.

Another prospective approach to identifying monogenean species employs molecular express diagnostic methods (Fromm et al., 2013, 2014; Mieszkowska et al., 2018; Vodiasova et al., 2022). However, it requires specialised equipment that is rarely available on fish farms. Therefore, microscopic examination and express diagnostic methods complement one another rather than replace each other.

From this perspective, machine learning methods in aquaculture present a promising approach for developing practical tools that can be directly used by farmers on fish farms. The advantage of these applications is that they require minimal equipment; however, this often results in reduced diagnostic accuracy compared to traditional laboratory methods. Sun and Li (2016) offered an expert system for diagnosing fish diseases, implemented as an Android application. It is an extensive database of fish diseases and their treatment methods. There are two types of users: farmers and experts. Experts keep the contents of the database up to date, and farmers are able to diagnose diseases and identify treatments, which match a given list of symptoms. Lou et al. (2007) developed an automated diagnostic system for the aquacultured shrimp *Penaeus vannamei*. The system comprises a database of possible diseases, where each entry enhanced by example photographs and a summary of relevant information. The search algorithm allows the user to search for a diagnosis by using either a set of symptoms or photographs. Lopes et al. (2011) proposed a neural network model for diagnosing fish diseases caused by protozoa and bacteria using a given set of Boolean indicators.

In this paper, we present a neural network algorithm that makes it possible to accurately identify members of the order Dactylogyridea in photographs taken through the ocular lens of the light microscope with a smartphone camera. This specific order is a valuable model object due to the availability of biological material. Moreover, it holds practical significance because it includes several members that are recognised as causative agents of diseases commonly associated with fish mortality. Our aim was to produce a tool that could be used directly on fish farms. We propose slightly stricter requirements for equipment compared to other similar studies assuming that at least a light microscope and a smartphone are available. However, these tools are still insufficient for producing images of comparable quality to those used for developing algorithms for monogenean species discrimination or for automated systems design. Since we aimed at a more general recognition level (between the orders), lower-quality images were reasonably good. It is commonly accepted that artificial neural networks are capable of learning meaningful features from low-resolution photographs of fish body surfaces for detecting disease symptoms. However, for distinguishing monogenean species, these models were typically applied to a pre-selected set of meaningful features manually measured in high-quality images. Our results show that an artificial neural network was able to successfully learn the distinctive features of the order Dactylogyridea from microscope images. The recognition accuracy was comparable with the results obtained by other studies that employed neural networks or classical machine learning methods (KNN, PCA, and others).

We trained our neural network algorithm on images with manually removed backgrounds. This means that potential users will need to perform the same background removal procedure on their data before using the algorithm. While this adds extra effort for the user, we believe it will help avoid false classifications caused by background noise. Additionally, given that convolutional neural networks are effective for image segmentation tasks, one could train a separate model for background removal. The results from this model would need to be verified and corrected by the user as necessary. We plan to consider this task in our future work.

Our approach clearly faces the same challenges as all artificial neural networks: it strongly depends on the amount and quality of training data. While we can maintain a reasonable minimum amount of training data through the proper application of augmentation methods, the labor-intensive collection of new biological data will be necessary to include additional orders of monogeneans. Nevertheless, the biological material used in this study can be integrated into future efforts to address these tasks. Further efforts would involve examining other orders, some members of which pose significant threats to the health of fish in aquaculture and natural water bodies in southern Russia (for example, Gyrodactylidea, Mazocraeidea or Diclybothriidea). This endeavor will require additional data, along with optimisation of the neural network architecture and image augmentation algorithms. As there are approximately seven important orders of fish parasites in the south of Russia, collection of the required amount of biological material seems to be a feasible task. The development of neural network classifiers capable of identifying these important orders of parasites could serve as a valuable resource for novice specialists. This automated system can be a tool to optimise routine work and reduce incorrect identifications and wrong treatment decisions.

Another avenue for advancing this research is the development of neural network algorithms for more precise identification of monogeneans at the family and genus levels. Addressing this issue has notable practical implications, since certain species within the order Dactylogyridea (*Dactylogyrus vastator*, *D. extensus*, etc.) can cause mass mortality in fish populations. To achieve more accurate classification, it is essential to ensure that digital photographs of parasites accurately depict critical taxonomic features of monogeneans, including the chitinoid structures of the attachment disc and reproductive system. Consequently, a different organisation of training data will be necessary, and there may also be additional requirements regarding the magnification power of the microscope and photographic resolution. These matters warrant further investigation.

Conclusions

This paper presents a neural network algorithm capable of accurately identifying members of the order Dactylogyridea in digital photographs taken through the ocular lens of the light microscope with a smartphone camera. The lowered equipment requirements make our method suitable for use on fish farms. The results obtained in this study hold practical value as many members of this order are pathogens causing dangerous diseases that lead to fish mortality. After being extended to other orders of parasites, the proposed method could simplify the work of novice specialists and to contribute to faster diagnosis of diseases, as well as the development of prevention and treatment methods.

References

Ahmed M. S., Aurpa T. T., Azad M. A. K. (2022) Fish disease detection using image based machine learning technique in aquaculture. *Journal of King Saud University – Computer and Information Sciences*, 34(8): 5170–5182

Ali R., Hussain A., Bron J. E., Shinn A. P. (2011) Multi-stage classification of Gyrodactylus species using machine learning and feature selection techniques. *2011 11th International Conference on Intelligent Systems Design and Applications*. New York, IEEE, p. 457–462

Ali R., Hussain A., Bron J. E., Shinn A. P. (2012) The use of ASM feature extraction and machine learning for the discrimination of members of the fish ectoparasite genus Gyrodactylus. *Neural Information Processing. ICONIP 2012. Lecture Notes in Computer Science, Vol 7666*. Huang T., Zeng Z., Li C., Leung C. S. (Eds.) Springer, Berlin, Heidelberg, p. 256–263

Awalludin E. A., Arsad T. N. T., Hj Wan Yussof W. N. J. (2020) A review on image processing techniques for fisheries application. *Journal of Physics: Conference Series*, 1529: 052031

Bauer O. N., Musselius V. A., Strelkov Yu. A. (1981) *Pond fish diseases*. Moscow, Legkaya i pishchevaya promyshlennost', 320 p. (in Russian)

Bauer O. N. (Ed.) (1985) *Key to parasites of freshwater fishes in the fauna of the USSR: Vol. 2. Parasitic multicellular organisms (first part)*. Leningrad, Nauka, 425 p. (in Russian)

Butko M. P., Popov P. A., Lemyaseva S. V., Onishchenko D. A. (2017) Fish invasive diseases dangerous for humans and animals. *Russian Journal "Problems of Veterinary Sanitation, Hygiene and Ecology"* [Rossiiskii zhurnal «Problemy veterinarnoi sanitarii, gigieny i ekologii»], 4: 121–127 (in Russian)

Bykhovskaya-Pavlovskaya I. E. (1985) *Parasitological study of fish*. Leningrad, Nauka, Leningrad Branch, 109 p. (in Russian)

FAO (2022) *The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation*. Rome, FAO. https://doi.org/10.4060/cc0461en

Feist S. W., Thrush M. A., Dunn P., Bateman K., Peeler E. J. (2019) The aquatic animal pandemic crisis. *Revue Scientifique et Technique*, 38(2): 437–457

Fromm B., Worren M. M., Hahn C., Hovig E., Bachmann L. (2013) Substantial loss of conserved and gain of novel MicroRNA families in flatworms. *Molecular Biology and Evolution*, 30(12): 2619–2628

Fromm B., Burow S., Hahn C., Bachmann L. (2014) MicroRNA loci support conspecificity of Gyrodactylus salaris and Gyrodactylus thymalli (Platyhelminthes: Monogenea). *International Journal for Parasitology*, 44(11): 787–793

Gaevskaya A. V. (2004) *Parasites and diseases of marine and oceanic fishes in natural and culture conditions*. Sevastopol, EKOSI- Gidrofizika, 237 p. (in Russian)

Goodfellow I., Bengio Y., Courville A. (2016) *Deep learning*. Cambridge, MA, MIT Press, 800 p.

Gusev A. V. (1983) *Methodology for collection and preparation of fish parasite monogenean samples*. Leningrad, Nauka, 48 p. (in Russian)

Hasan N., Ibrahim S., Aqilah Azlan A. (2022) Fish diseases detection using convolutional neural network (CNN). *International Journal of Nonlinear Analysis and Applications*, 13(1): 1977–1984

Ioffe S., Szegedy C. (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv:1502.03167*

Isukapalli R., Elgammal A., Greiner R. (2006) Learning to detect objects of many classes using binary classifiers. *Computer Vision – ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, Vol 3951*. Leonardis A., Bischof H., Pinz A. (Eds.) Springer, Berlin, Heidelberg, p. 352–364

Krizhevsky A., Sutskever I., Hinton G. E. (2017) ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 84–90

Li D., Li X., Wang Q., Hao Y. (2022) Advanced techniques for the intelligent diagnosis of fish diseases: a review. *Animals*, 12(21): 2938

Li J., Lian Z., Wu Z., Zeng L., Mu L., Yuan Y., Bai H., Guo Z., Mai K., Tu X., Ye J. (2023) Artificial intelligence–based method for the rapid detection of fish parasites (Ichthyophthirius multifiliis, Gyrodactylus kobayashii, and Argulus japonicus). *Aquaculture*, 563(Part 1): 738790

Lopes J. N. S., Goncalves A. N. A., Fujimoto R. Y., Carvalho J. C. C. (2011) Diagnosis of fish diseases using artificial neural networks. *International Journal of Computer Science Issues*, 8(6): 68–74

Lou D., Chen M., Ye J. (2007) Study on a fish disease case reasoning system based on image retrieval. *New Zealand Journal of Agricultural Research*, 50(5): 887–893

Mieszkowska A., Górniak M., Jurczak-Kurek A., Ziętara M. S. (2018) Revision of Gyrodactylus salaris phylogeny inspired by new evidence for Eemian crossing between lineages living on grayling in Baltic and White sea basins. *PeerJ*, 6: e5167

Nikolenko S. I., Kadurin A., Arkhangelskaya E. V. (2018) *Deep learning. A dive into the world of neural networks*. Saint Petersburg, Piter, 481 p. (in Russian)

Park J.-S., Oh M.-J., Han S. (2007) Fish disease diagnosis system based on image processing of pathogens' microscopic images. *2007 Frontiers in the Convergence of Bioscience and Information Technologies*. New York, IEEE, p. 878–883

Pauzi S. N., Hassan M. G., Yusoff N., Harun N. H., Abu Bakar A. H., Kua B. C. (2021) A review on image processing for fish disease detection. *Journal of Physics: Conference Series*, 1997: 012042

Pugachev O. N., Gerasev P. I., Gussev A. V., Ergens R., Khotenowsky I. (2009) *Guide to Monogenoidea of freshwater fish of Palaearctic and Amur regions*. Milan, Ledizioni-LediPublishing, 564 p.

Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A. C., Fei-Fei L. (2015) ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3): 211–252

Shorten C., Khoshgoftaar T. M. (2019) A survey on image data augmentation for deep learning. *Journal of Big Data*, 6: 60

Simonyan K., Zisserman A. (2015) Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*

Siradjuddin I. A., Subroto A., Muntasa A. (2020) Binary classification using convolutional neural network for attributes identification of pedestrian image. *Journal of Physics: Conference Series*, 1569: 022058

Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. (2014) Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15: 1929–1958

Strona G., Montano S., Seveso D., Galli P., Fattorini S. (2014) Identification of Monogenea made easier: a new statistical procedure for an automatic selection of diagnostic linear measurements in closely related species. *Journal of Zoological Systematics and Evolutionary Research*, 52(2): 95–99

Sun M., Li D. (2016) Aquatic animal disease diagnosis system based on android. *Computer and Computing Technologies in Agriculture IX. CCTA 2015. IFIP Advances in Information and Communication Technology, Vol 478*. Li D., Li Z. (Eds.) Springer, Cham, p. 115–124

Vignon M. (2011a) Putting in shape: towards a unified approach for the taxonomic description of monogenean haptoral hard parts. *Systematic Parasitology*, 79(3): 161–174

Vignon M. (2011b) Inference in morphological taxonomy using collinear data and small sample sizes: Monogenean sclerites (Platyhelminthes) as a case study. *Zoologica Scripta*, 40(3): 306–316

Viola P., Jones M. (2001) Fast and robust classification using asymmetric AdaBoost and a detector cascade. *Proceedings of the 15th International Conference on Neural Information Processing Systems: Natural and Synthetic (NIPS'01)*. MIT Press, Cambridge, MA, USA, 1311–1318

Vodiasova E. A., Chelebieva E. S., Shikhat O. V., Atopkin D. M., Dmitrieva E. V. (2022) Molecular-genetic approaches to species identification of platyhelminthes of the genus Ligophorus (Monogenea) parasitising flathead mullet. *Vavilov Journal of Genetics and Breeding* [Vavilovskii zhurnal genetiki i selektsii], 26(3): 290–297 (in Russian)

Waleed A., Medhat H., Esmail M., Osama K., Samy R., Ghanim T. M. (2019) Automatic recognition of fish diseases in fish farms. *2019 14th International Conference on Computer Engineering and Systems (ICCES)*. New York, IEEE, p. 201–206

World Register of Marine Species. Dactylogyridea [Electronic resource]: https://www.marinespecies.org/aphia.php?p=taxdetails&id=517897 (accessed 24.05.2023).

Yang X., Zhang S., Liu J., Gao Q., Dong S., Zhou C. (2021) Deep learning for smart fish farming: applications, opportunities and challenges. *Reviews in Aquaculture*, 13(1): 66–90

Zhan Y., Geng X., Liu Y., Zhang Y. (2020) Research on the classification of fish parasitoids based on digital image processing technology. *IOP Conference Series: Materials Science and Engineering*, 768: 072108

Yousef Kalafi E., Tan W.B., Town C., Dhillon S.K. (2016) Automated identification of Monogeneans using digital image processing and K-nearest neighbour approaches. *BMC Bioinformatics*, 17(Suppl. 19): 511