Short communication

Identification of Baikal phytoplankton inferred from computer vision methods and machine learning

Lysenko A.V.[1,2]*, Oznobikhin M.S.[1,2], Kireev E.A.[1,2], Dubrova K.S.[1], Vorobyeva S.S.[1]

[1] *Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya str., 3, Irkutsk, 664033, Russia*
[2] *Institute of Mathematics and Information Technologies, Irkutsk State University, Gagarina str., 20, Irkutsk, 664003, Russia*

ABSTRACT. This study discusses the problem of phytoplankton classification using computer vision methods and convolutional neural networks. We created a system for automatic object recognition consisting of two parts: analysis and primary processing of phytoplankton images and development of the neural network based on the obtained information about the images. We developed software that can detect particular objects in images from a light microscope. We trained a convolutional neural network in transfer learning and determined optimal parameters of this neural network and the optimal size of using dataset. To increase accuracy for these groups of classes, we created three neural networks with the same structure. The obtained accuracy in the classification of Baikal phytoplankton by these neural networks was up to 80%.

1. Introduction

In the modern world, machine-learning methods are becoming increasingly applicable in different fields of science. For example, they are used in the sequence data mining of DNA (Yang et al., 2020). Computer vision methods and neural networks are widely used in biology for object recognition (Majaj and Pelli, 2018; Sánchez et al., 2019). In biology, recognition of objects in an image provides automatization of statistical analysis of organisms, biomass counting, etc. Usually, biologists analyze these images manually, whereas using special software will help to classify objects faster. Additionally, a software interface will allow the user to use the program for one specific image or for a set of images. It makes the program more visually clear for any user.

One of the most widely used image classification methods is convolutional neural networks. Convolutional neural networks are a very large class of architectures, the main idea of which is to alternate convolutional layers and pooling layers. The structure of this network is unidirectional and multilayered. A convolutional neural network consists of an input layer, hidden layers and an output layer. These hidden layers usually consist of convolutional layers, pooling layers, fully connected layers, and normalization layers (Hussain et al., 2018). Convolution neural networks have applications (Yamashita et al., 2018) in image classification, image segmentation, object detection, image and video recognition as well as in natural language processing. Neural networks that are used for classification usually contain the same number of output neurons in the last layer as the number of classes. This output vector is a set of probabilities of images belonging to each class. The highest probability value corresponds to the neural network prediction.

Transfer learning is often used to improve the accuracy of the prediction of the neural network. Transfer learning is a machine learning technique that focuses on using the experience gained from solving one problem to solve another similar problem. At first, the neural network is trained on a large amount of data, then on the purpose dataset. This method is often used when a purpose dataset is smaller than the original dataset used to train the pre-trained model (Larsen-Freeman, 2013). Also, transfer learning is applied in such types of biology tasks as plant image classification (Tapas, 2016).

There are many computer vision methods for image segmentation and object detection. It can be also done using neural networks. The application of computer vision methods for segmentation requires converting an image to grayscale and then binarizing it. One of the most popular binarization techniques is Otsu's method. This algorithm returns a single intensity threshold that separates pixels into two classes, foreground and background (Otsu, 1979). The Otsu's method also

*Corresponding author.
E-mail address: allessouth@gmail.com (A.V. Lysenko)

has an improved version to support multiple image levels, which is called the multi-Otsu method (Liao et al., 2001). Another popular binarization technique is Canny Edge Detection (1986). This technique allows the detection of a wide range of edges in images. For our dataset, Otsu's method demonstrated the best results. After binarization of the image, the pixels that belong to the regions of interest take the value of 1.

Data augmentation is a technique for creating additional training data from existing data. It acts as a regularizer and helps reduce overfitting when training a machine learning model (Shorten and Khoshgoftaar, 2019). To achieve good results, neural networks must be trained on a large amount of data. Therefore, if the original dataset contains a limited number of images, it is necessary to perform data augmentation to improve the results of the work of the neural network. There are many techniques to do data augmentation such as the popular horizontal flipping, vertical flipping, random crop, random rotation, etcetera (Shorten and Khoshgoftaar, 2019). Also, there are color augmentation techniques, and the most popular of them are histogram equalization, white-balancing, sharpening, blurring, and enhancing contrast or brightness (Galdran et al., 2019). Several of these techniques can be combined. Implemented augmentation techniques increase the original dataset by a factor of 10.

In our study, we try to connect computer vision methods and machine learning for the classification of phytoplankton in Lake Baikal.

2. Materials and methods

Phytoplankton samples were collected from the shallow zone at the Listvyanka settlement (southern basin of Lake Baikal) 51.868022 N, 104.82959 E from February to May 2021. This area has no continuous ice cover during winter. The samples were fixed with the Utermöhl solution and concentrated by sedimentation. The concentrate was placed into a 0.1 mL cell and examined under an Amplival (Carl Zeiss, Germany) microscope at 640x magnification.

The original dataset consisted of 2705 images with dimensions of 3488x2616, each of which might contain several instances of each class of recognizable objects. 4622 original phytoplankton images were obtained by manual segmentation. Our neural network was constructed based on the phytoplankton taxa *Cyclotella minuta* Antipova (162 images), *Aulacoseira baicalensis* (K. Meyer) Simonsen (295 images), *Synedra acus* subsp. *radians* (Kütz.) Skabitsch (213 images), *Cryptomonas* sp. (527 images), *Rhodomonas pusilla* (Bachm.) Javorn. (155 images), *Dynobrion cylindricum* Imhof (411 images), *Gyrodinium helveticum* (Penard) (226 images), *Peridinium baicalense* Kiselev et Cvetkov (314 images), *Gymnodinium baicalense* (197 images), and *Koliella longiseta* (Vischer) Hindak (205 images).

The Watershed algorithm was applied to image segmentation. The input image was blurred with mean shift filtering and converted to grayscale. The resulting image was binarized with Otsu's method. Then, the distance to the closest zero pixels for each pixel of the source image was calculated. After distance calculating, we found local maxima and implemented the Watershed algorithm. The result of applying this algorithm is a matrix of labels that can be converted into an image. Segmented objects have different numeric labels, and it allows selecting them according to these numerical labels.

We used a pre-trained Xception neural network (Chollet, 2017) as the initial network and created a 'target' neural network using transfer learning (Pratt and Jennings, 1996) technology, in which two dense layers were used, including 1024 and 512 neurons. We used the following parameters of the neural network: the rectified linear unit (ReLU) activation function, dropout of 30% (to avoid retraining of the neural network) and the output layer, with softmax activation function. To train the neural network, the following datasets were used for each of the 10 species: 184 images for training and 30 images for validation. Learning took place for 100 epochs with the stochastic gradient descent (SGD) optimization function and the learning rate parameter equal to 0.005.

3. Results and discussion

The analyzed phytoplankton taxa were different in i) outlet contours from oval to rectangular (Fig. 1) and ii) grouping of objects from single to mass clusters (Fig. 2). As a result, the program detects single objects in the images with an accuracy of 90%. In images with noisy regions or other objects (such as incomplete parts of objects), the accuracy decreased to 72%. In the case of objects intersection when many objects have a common border the accuracy reaches 59%. For a sample of images of each type, the accuracy was 62% (Fig. 2).

The trained neural network most accurately identifies *A. baicalensis*, *Cryptomonas* sp., *C. minuta*, *D. cylindricum*, *K. longiseta*, *P. baicalense*, and *S. acus* with an accuracy > 80% (Table). The most mistakes concerned the identification of *Rh. pusilla* with an accuracy of 48% where 48% of images were regarded as *Cryptomonas* sp. It was unexpected because *Cryptomonas* sp. was detected with high reliability (Table) and confused only with *Rh. pusilla* in 11% cases. For this reason, the special train neural network was formed for only these two species. The accuracy of this neural network is 96% for the recognition of *Cryptomonas* sp. and 77% for *Rh. pusilla*. Moreover, other spatial network was formed for the assemblage of *G. helveticum*, *P. baicalense* and *G. baicalense*. The accuracy of this neural network was 90% for *G. helveticum*, 88% for *G. baicalense*. and 96% for *P. baicalense*. The architecture of these neural networks was the same; the differences were only in the number of neurons in the output layer.

At present, the phytoplankton assemblage of Lake Baikal is characterized by 70 taxa and their species. The amount of phytoplankton changed from 0.015 to 136 million cells L^{-1} with biomass of 0.004-2.13 g m^{-3}. Diatoms (20 taxa) were represented by *A. baicalensis*, *A. islandica*, *C. minuta*, *C. baicalensis*, *Nitzschia graciliformis*, *S. acus*, and *Stephanodiscus meyeri*. *Gyrodinium*
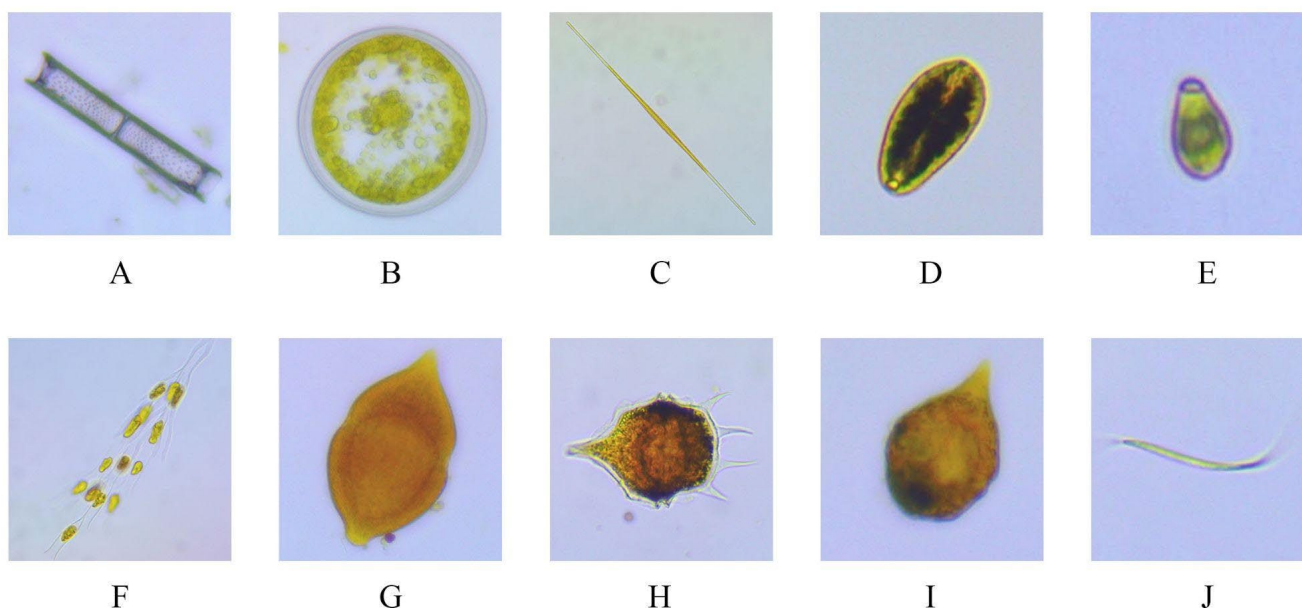
Fig.1. Sample images of phytoplankton taxa that can be recognized by software in this study (A - *Aulacoseira baicalensis*, B - *Cyclotella minuta*, C - *Synedra acus*, D - *Cryptomonas*, E - *Rhodomonas pusilla*, F - *Dynobrion cylindricum*, G - *Gyrodinium helveticum*, H - *Peridinium baicalense*, I - *Gymnodinium baicalense*, and J - *Koliella longiseta*).

helveticum, G. baicalense, P.baicalense and *Glenodinium* sp. predominate in Dinophyta. *Rhodomonas pusilla* (up to 6064 thousand cells L^{-1}) dominated Cryptophyta. Chrysophyta were formed by *Chrysochromulina parva* (up to 6439 thousand cells L-1) and *D.cylindricum* (up to 192 thousand cells L-1). Chlorophyta was represented by *Monoraphidium contortum, M. arcuatum, K. longiseta,* and *Chlamydomonas* sp (Vorobyeva, 2018; Bondarenko et al., 2020). Hence, the constructed three networks can help to identify the main taxon of the Baikal phytoplankton with high accuracy.

4. Conclusions

In this study, we tested the neural networks for the identification of the Baikal phytoplankton assemblage. One neural network did not define accurately large data set of species of Diatoms, Dinophyta, Cryptophyta, Chrysophyta, and Chlorophyta in the Baikal phytoplankton. We constructed three networks for i) *A. baicalensis, C. minuta, D. cylindricum, K.longiseta, P.*

baicalense, S. acus; ii) *Cryptomonas* sp. and *Rh. usilla;* iii) *G. helveticum, P. baicalense* and *G. baicalense.* The accuracy of these neural networks was >80%. The software developed for object detection allows segmenting of many objects in images. This software can detect isolated objects with high accuracy. This is more difficult to allocate objects in noisy images than to allocate single objects in non-noisy images. In the future, it will be necessary to supplement the dataset with rare species because the number of images, on which the neural network is trained, in some species is more than 50% of the number of images of this species in the general dataset that we tested. Therefore, the results are undoubtedly less reliable.

Acknowledgments
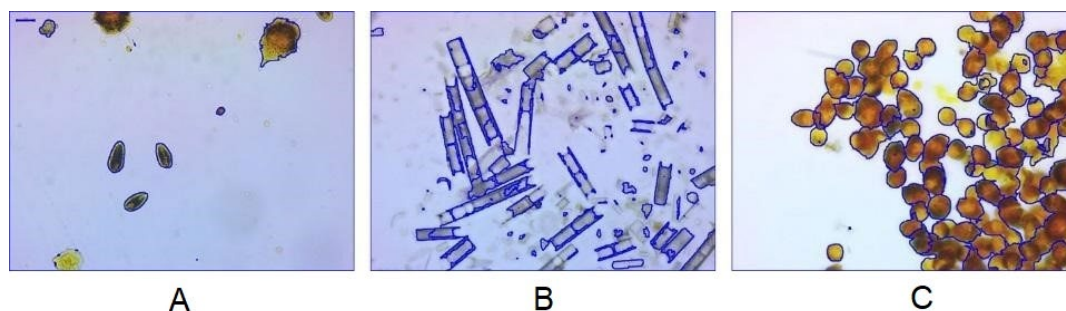
Fig.2. Sample images of phytoplankton clusters with contour detection examples (A - isolated objects, B - objects in noisy image, C - intersecting objects in image without noise).

Table. Recognition results.

actual class in test dataset	distribution of predictions, percent									
	A. baicalensis	*Cryptomonas* sp.	*C. minuta*	*D. cylindricum*	*G. helveticum*	*G. baicalense*	*K. longiseta*	*P. baicalense*	*Rh. pusilla*	*S. acus*
A. baicalensis	**94**	0	0	0	0	0	0	0	0	5
Cryptomonas sp.	0	**89**	0	0	0	0	0	0	11	0
C. minuta	0	0	**98**	0	0	1	0	0	0	0
D. cylindricum	0	0	0	**100**	0	0	0	0	0	0
G. helveticum	0	4	0	0	**64**	13	0	17	1	0
G. baicalense.	0	5	0	0	9	**68**	0	17	1	0
K. longiseta	0	0	0	1	0	0	**95**	0	0	3
P. baicalense	0	1	0	1	3	6	0	**89**	0	0
Rh. pusilla	0	48	1	0	0	0	0	1	**48**	0
S. acus	5	0	0	2	0	0	8	0	0	**84**

In this table, the rows correspond to the actual value, and the columns - to the predicted value; when they intersect, the prediction is considered correct.

Conflict of interests

The authors declare no conflicts of interest.

References

Canny J. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6): 679-698. DOI: 10.1109/TPAMI.1986.4767851

Chollet F. 2017. Xception: deep learning with depthwise separable convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition. DOI: 10.1109/CVPR.2017.195

Galdran A., Alvarez-Gila A., Meyer M.I. et al. 2017. Data-driven color augmentation techniques for deep skin image analysis. arXiv:1703.03702v1.

Hussain M., Bird J.J., Faria D.R. 2018. A study on CNN transfer learning for image classification. In: 18th Annual UK Workshop on Computational Intelligence. DOI: 10.1007/978-3-319-97982-3 16

Larsen-Freeman D. 2013. Transfer of learning transformed. Language Learning 63(s1). DOI: 10.1111/j.1467-9922.2012.00740.x

Liao P.-S., Chen T.-S., Chung P.-C. 2001. A fast algorithm for multilevel thresholding. Journal of Information Science and Engineering 17(5): 713-727. DOI: 10.6688/JISE.2001.17.5.1

Otsu N. 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1): 62-66. DOI: 10.1109/TSMC.1979.4310076

Pratt L., Jennings B. 1996. A survey of transfer between connectionist networks. Connection Science 8(2): 163-184. DOI: 10.1080/095400996116866

Sánchez C., Cristóbal G., Bueno G. 2019. Diatom identification including life cycle stages through morphological and texture descriptors. PeerJ 7. DOI: 10.7717/peerj.6770

Shorten C., Khoshgoftaar T.M. 2019. A survey on image data augmentation for deep learning. Journal of Big Data 6(1). DOI: 10.1186/s40537-019-0197-0

Tapas A. 2016. Transfer learning for image classification and plant phenotyping. In: Second International Conference on Electronics, Communication and Aerospace Technology. DOI: 10.1109/ICECA.2018.8474802

Vorobyeva S.S. 2018. Phytoplankton assemblages of the Southern Baikal in 1990-1995 and 2016-2018. Limnology and Freshwater Biology. 1(2): 141-143. DOI: 10.31951/2658-3518-2018-A-2-141

Bondarenko N.A., Vorobyova S.S., Zhuchenko N.A. et al. 2020. Current state of phytoplankton in the littoral area of Lake Baikal, spring 2017. Journal of Great Lakes Research 46(1): 17-28. DOI: 10.1016/j.jglr.2019.10.001

Yang A., Zhang W., Wang J. et al. 2020. Review on the application of machine learning algorithms in the sequence data mining of DNA. Frontiers in Bioengineering and Biotechnology 8. DOI: 10.3389/fbioe.2020.01032

Yamashita R., Nishio M., Do R.K.G. et al. 2018. Convolutional neural networks: an overview and application in radiology. Insights into Imaging 9(4). DOI: 10.1007/s13244-018-0639-9

Majaj N.J., Pelli D.G. 2018. Deep learning - using machine learning to study biological vision. Journal of Vision 18(13):2: 1-13. DOI: 10.1167/18.13.2