Scientific Journals
of the Maritime University of Szczecin

Zeszyty Naukowe
Akademii Morskiej w Szczecinie

2021, 67 (139),
ISSN 1733-8670 (Printed)
ISSN 2392-0378 (Online)
Article accepted for publication

Received: 24.03.2021
Accepted: 08.09.2021
Published: 30.09.2021

Selection of an artificial pre-training neural network for the classification of inland vessels based on their images

Katarzyna Bobkowska[1], Izabela Bodus-Olkowska[2]

https://orcid.org/0000-0003-4968-3407
https://orcid.org/0000-0003-4366-0116

[1] Marine Technology Ltd.
4 lok. 6 Roszczynialskiego St., 81-521 Gdynia, Poland
e-mail: k.bobkowska@marinetechnology.pl
Gdańsk University of Technology, Faculty of Civil and Environmental Engineering
11/12 Gabriela Narutowicza St., 80-233 Gdańsk,Poland
[2] Maritime University of Szczecin
1-2 Wały Chrobrego St., 70-500 Szczecin, Poland
e-mail: i.olkowska@am.szczecin.pl
corresponding author

Keywords: image classification, inland vessels, ANN, pre-trained neural network, GoogLeNet, AlexNet, SqeeezeNet

JEL Classification: R49, R490, R420, R410, Z00, Z000

Abstract

Artificial neural networks (ANN) are the most commonly used algorithms for image classification problems. An image classifier takes an image or video as input and classifies it into one of the possible categories that it was trained to identify. They are applied in various areas such as security, defense, healthcare, biology, forensics, communication, etc. There is no need to create one's own ANN because there are several pre-trained networks already available. The aim of the SHREC projects (automatic ship recognition and identification) is to classify and identify the vessels based on images obtained from closed-circuit television (CCTV) cameras. For this purpose, a dataset of vessel images was collected during 2018, 2019, and 2020 video measurement campaigns. The authors of this article used three pre-trained neural networks, GoogLeNet, AlexNet, and SqeezeNet, to examine the classification possibility and assess its quality. About 8000 vessel images were used, which were categorized into seven categories: barge, special-purpose service ships, motor yachts with a motorboat, passenger ships, sailing yachts, kayaks, and others. A comparison of the results using neural networks to classify floating inland units is presented.

Introduction

Artificial neural networks are currently used in many fields, including medicine, environmental assessments, transport calculations, and situational assessments based on photos. Neural networks are an excellent tool for the automatic detection and classification of items shown in images. The use of a neural network as a dedicated analysis tool significantly facilitates such work. Creating a dedicated neural network requires IT skills, and overtraining requires a large set of learning data (Wlodarczyk-Sielicka & Polap, 2019; Polap & Wlodarczyk-Sielicka, 2020). In some cases, it is not possible to collect a sufficiently large number of images for analysis. In this case, pre-trained neural networks can be used. Matlab software offers several such networks, including VGG-16, VGG-19, Xception, AlexNet, GoogLeNet, Inception, and SqueezeNet. These are networks with various parameters that have been developed and trained in several dozen categories and have been made available for research purposes in various scientific fields. In the scientific literature, there are many studies based on the use of these networks. They mainly concern the analysis of medical images but also forensics, object monitoring, the detection of people based on a photo of their face, observation of cargo, or classification of plant diseases, and even for analysing the seafloor and identifying underwater objects. Pre-trained neural networks are readily available, ready-to-use, and their training process is relatively fast and easy. They provide sufficient results when used for image classification, which has been presented in many research papers (Espinosa, Velastin & Branch, 2017; Salavati & Mohammadi, 2018; Sang-Geol et al., 2018; Alaskar et al., 2019; Khan, Zhang & Kumar, 2019; Konovalov et al., 2019; Lee et al., 2019; Wang et al., 2019; Hassanpour & Malek, 2020). Since building a new neural network dedicated to a given task requires time and skills, it is much easier to use already-available, over-trained networks for cognitive purposes. Thanks to this, researchers who do not regularly deal with neural networks, i.e., non-IT specialists, can easily adapt pre-trained networks to their research fields.

The task of classifying floating objects based on photos from video cameras is one of the goals of the Automatic Ship Recognition and Identification (SHREC) project. CCTV monitoring consists of many cameras located at strategic points, i.e., on bridges and in marinas. One such system covers the port of Szczecin–Świnoujście in northwest Poland. One of the goals of the project is to classify vessels; therefore, several pre-trained neural networks were tested, and the one that gave the best results was selected. The project is funded by the Polish LIDER NCBiR program (Wawrzyniak & Stateczny, 2018; Bobkowska & Wawrzyniak, 2019). The aim of this article is to compare the most popular neural networks for classifying conventional ships.

Characteristics of the input dataset

The main action in the implementation of the task was to develop a class of vessels in the surveyed area. The decision was made based on the observation of selected locations near bridges and marina aquatics. As the project involves the identification of conventional (STCW) and non-conventional units, 21 classes were initially established (Bobkowska & Bodus-Olkowska, 2020):

1. Kayak, Pedalo, Rowing Boat
2. Small Boat, Motor Boat
3. Motor Yacht
4. Sailing Yacht with a mast
5. Sailing Yacht with a mast down
6. Large Motor Yacht
7. Sailing Ship
8. Barge
9. Inland Pusher
10. Pushed Convoy
11. Water Services
12. Small Vessel
13. Medium Vessel
14. Large Vessel
15. Navy Vessel
16. Special Vessel
17. Passenger Vessel
18. Special-Purpose Service Ships
19. Fishing Vessel
20. Ships of Historical Value
21. Other

The measurement campaigns in 2018 and 2019 showed that the classes are too detailed. Due to the nature of the port, the investigation area does not have a sufficient number of different vessels in all of the proposed classes; therefore, the number of classes was reduced to seven:

1. Barges
2. Special-Purpose Service Ships
3. Motor Yachts and Motor Boats
4. Passenger Ships
5. Sailing Yachts
6. Kayaks
7. Other

STCW (International Convention on Standards of Training, Certification, and Watchkeeping) conventional ships (categories: small, medium, large), navy and fishing vessels, and ships with historical value were not taken into account in the further analysis of the article. Barges, inland pushers, and pushed convoys were combined into one class: barges. The special-purpose service ships class includes service units, hydrographic units, and port services. All small recreational craft, such as kayaks, pedalos, and rowing boats were added to the class “other”.

The input data set for testing included images of sample ships obtained from various websites and a set of photos selected from measurement campaigns carried out in 2018 and 2019. In total, the test dataset included more than 8000 photos of ships with a bow, stern, and a side view of the ship. 75% of the images were used for training purposes and 25% for validation.

Data and method

Selected pre-trained neural networks

There are many pre-trained neural networks available, and they are constantly being refined. Many of them were created as part of IT competitions or as a result of doctoral dissertations. For classification purposes, it is possible to integrate an original set of images into all pre-trained networks to obtain results. The networks used for research analysis in this article are described below.

GoogLeNet was proposed by research at Google. Its architecture consists of 22 layers, including pooling layers, some of which included a total of nine inception modules. The network trained on ImageNet (dataset available at https://image-net.org/) classifies images into 1000 object categories and on Places365 (dataset available at http://places2.csail.mit.edu/) into categories such as fields, parks, and runways. Both pre-trained networks have an image input size of 224 x 224, and also use a method called global average pooling at the end of the network analysis. This method decreases the number of trainable parameters to 0 and improves the top-1 accuracy by 0.6%. In GoogLeNet architecture, there is a fixed convolution size for each layer. In the Inception module 1×1, 3×3, 5×5 convolution, and 3×3 max-pooling are performed in parallel at the input, and the output of these are stacked together to generate the final output. Different-sized filters can better handle objects at multiple scales. GoogLeNet is widely used in medical aspects (Sang-Geol et al., 2018), object and obstacle detection (Salavati & Mohammadi, 2018), and malware detection (Khan, Zhang & Kumar, 2019).

AlexNet is a convolutional neural network that is eight layers deep: five convolutional layers and three fully-connected layers. AlexNet uses rectified linear units (ReLU) (Danqing, 2017) instead of the tanh function (hyperbolic tangent function). The greatest advantage is the training time. Multiple GPUs allow training by putting half of the model's neurons on one GPU and the other half on another GPU. This allows a bigger model to be trained in a much shorter training time. In AlexNet, the methodology of overlapping neurons is used, which reduces error by about 0.5%, and makes it harder to overfit. The overfitting problem is also reduced by two methods: data augmentation and dropout. The first one generates image translations and horizontal reflections, which increases the training set by a factor of 2048. AlexNet's authors also performed principal component analysis (PCA) on the RGB pixel values to change the intensities of RGB channels. The network has an image input size of 227 x 227. AlexNet is widely used in medical research (Wang et al., 2019), face recognition (Sang-Geol et al., 2018), vehicle detection (Espinosa, Velastin & Branch, 2017), biology (Konovalov et al., 2019), and many other fields (Alaskar et al., 2019).

SqueezeNet is a convolutional neural network that is 18 layers deep. SqueezeNet was developed by researchers at DeepScale, University of California, Berkeley and Stanford University and was released in 2016. The main goal was to develop a smaller neural network with fewer parameters. It was done by "squeezing parameters" – replacing 9X parameters: 3×3 filters with 1×1. Then, the "expanded" layers were convolution layers with a mix of 1×1 and 3×3 filters. Unlike other ANNs, where the activation map becomes smaller by the end of the network analysis, SqeezeNet has a larger activation map, which greatly increased the classification accuracy. Many uses of this network have been reported (Iandola et al., 2016; Lee et al., 2019; Hassanpour & Malek, 2020; Ucar & Korkmaz, 2020).

More specific information about pre-trained neural networks, their structure, the principle of operation, and training can be found in (Gallo, 2015; Basheer & Hajmeer, 2001).

Workflow procedure

For each of the three indicated artificial neural networks, a transfer learning process was carried out using Matlab 2020b with the Deep Learning Toolbox to adapt the artificial neural network to classify new classes. This process allowed for faster and easier learning of the network than learning from bases with random weights. The scheme of the procedure was:

1. Load Data
2. Load Pretrained Network
3. Replace Final Layers
4. Freeze Initial Layers
5. Train Network
6. Classify Validation Images

The entire procedure was based on the script provided by the Matlab software producer. Finally, the possibility of using the network was compared and it was assessed in terms of its use to classify ship images. The used criterion was the value of the classification quality in the form of a percentage of correctly classified images. Additionally, the quality of classification of individual classes was assessed.

Table 1. Number of images used to train and test artificial neural networks

Class	Number of images used in training	Number of images used in testing
Barge	434	80
Special-purpose service ships	1123	67
Motor yacht, with a motorboat	2560	376
Passenger ship	1654	210
Sailing yacht	389	237
Kayak	207	94
Other	698	10

Training and validation dataset

In the dataset for transfer learning, a total of 7065 images were used and 1074 for testing. The detailed specifications of the quantity are presented in Table 1.

Results and discussion

The implementation of the learning process was an easy procedure and was completed using a computer with the following parameters:

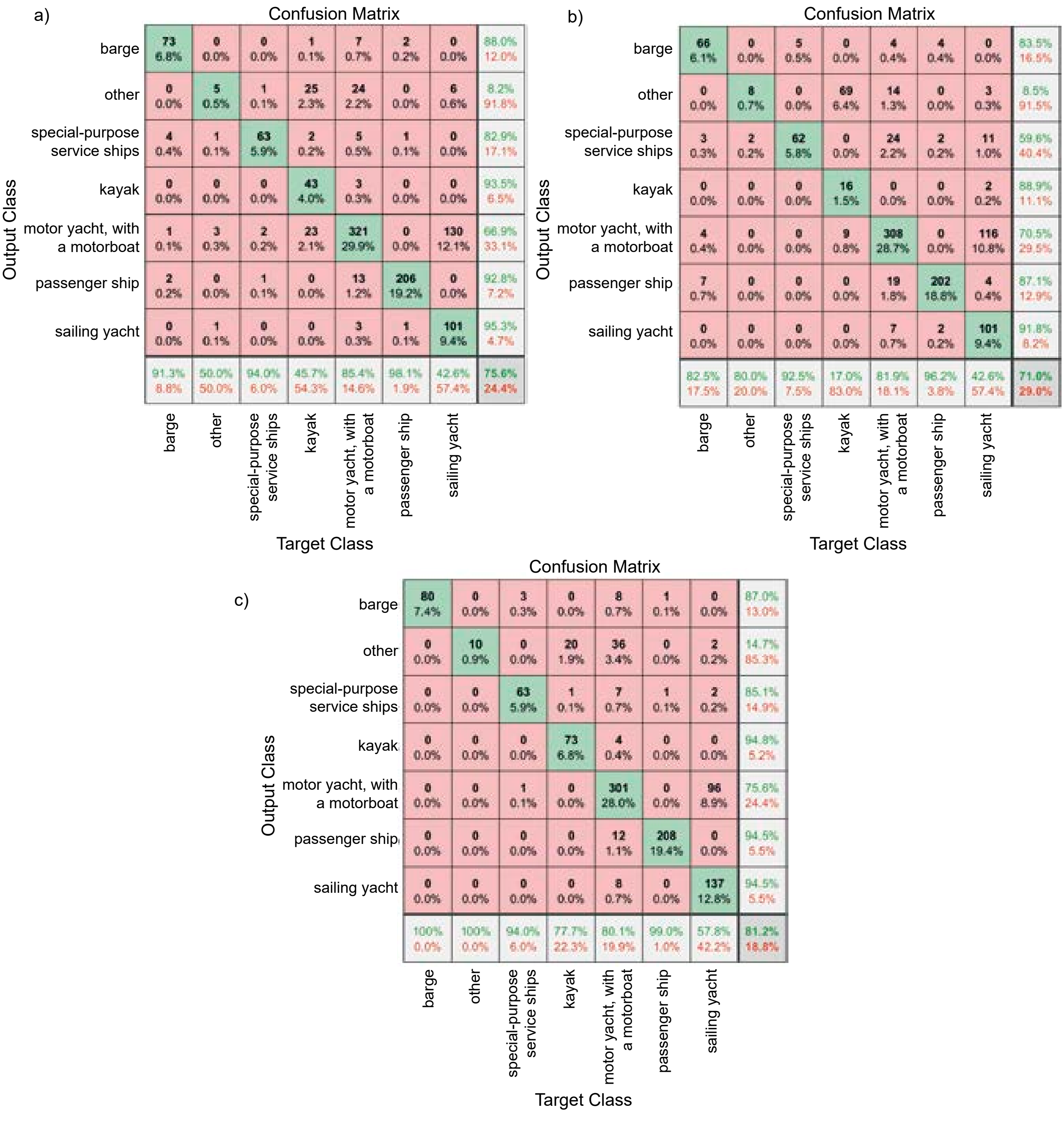

Figure 1. Confusion Matrix for a) SqueezNet, b) AlexNet, and c) GoogLeNet

- CPU: Intel® Xeon® CPU E3-1505M v5 @ 2.80 GHz,
- RAM 32.0 GB,
- 64-bit operating system,
- NVIDIA Quadro M1000M graphics card.

The times required to complete the learning task are presented in Table 2. The learning was carried out in each case in 6 epochs, after 635 iterations per epoch (3210 total).

Table 2. Time of the artificial neural network learning process

Artificial neural network	Elapsed time
GoogLeNet	29 min 35 sec
AlexNet	37 min 58 sec
SqueezNet	17 min 19 sec

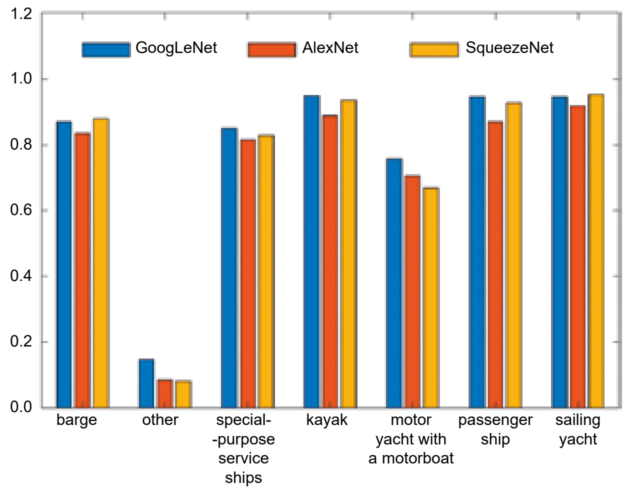

Figure 3. Output class

The detailed results in the form of a confusion matrix of the testing process for individual networks are presented in Figure 1.

To compare the classification quality of individual classes, the percentage results of the classification for specific networks are presented in Figure 2. In Figure 3, the results of the output class are presented, which show what percentage of images marked as a particular class was correctly layered as that class. For example, it means that the class "other" is classified with good quality (in the case of GoogLeNet and AlexNet) in Figure 2, while many images were also mistakenly classified as "other" in Figure 3.

To more accurately verify which classes were incorrectly classified, charts characterizing this phenomenon were prepared (Figure 4). The least problematic classes were barge, kayak, passenger ships, and sailing yachts. Nevertheless, in the cases of the classes of special-purpose service ships, motor yachts with motorboats, and other were confused and showed a repeated phenomenon regardless of the network. It should be noted, however, that for some networks, the level of misclassification was lower – it was the lowest for GoogLeNet. On the other hand, the other two performed comparably and worse. Based on the above charts, it can only be concluded that GoogLeNet coped best with its assumed task – the classification of unconventional vessels; however, it is possible to indicate classes for which new datasets – more numerous and more detailed – should be prepared to correctly classify them. An alternative is to search for a new artificial neural network that will better cope with this task to produce a higher-quality classification of the incorrectly classified units.

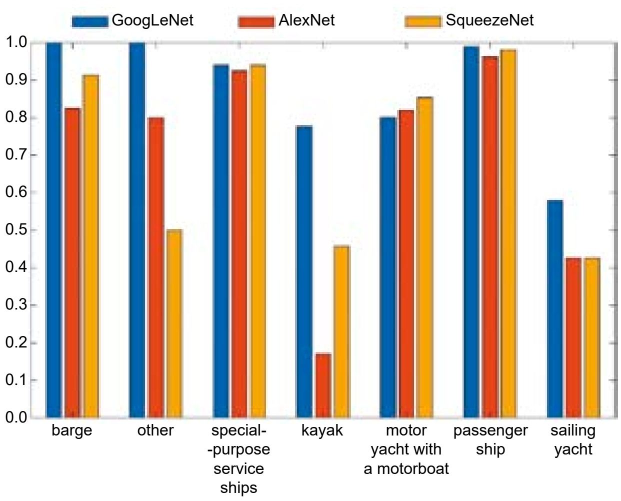

Figure 2. Target Class

Conclusions

The use of pre-trained neural networks is a convenient approach to adapt such networks to a specific task. The above considerations show that they can be used to classify ships. The learning time is short, and depending on the network, it was between 17 and 38 minutes. In the case of developing a new network, adding new values and weights is unattainable using a home computer. The best of the proposed options adjusting the GoogleNet network, which produced the highest-quality classification for the prepared training and testing datasets. Nevertheless, to obtain a highly satisfactory classification quality – close to 100 percent – errors related to the misclassification of other, motor yachts with a motorboat, special-purpose service ships, and sailing yachts (which were confused with a folded mast), should be eliminated.

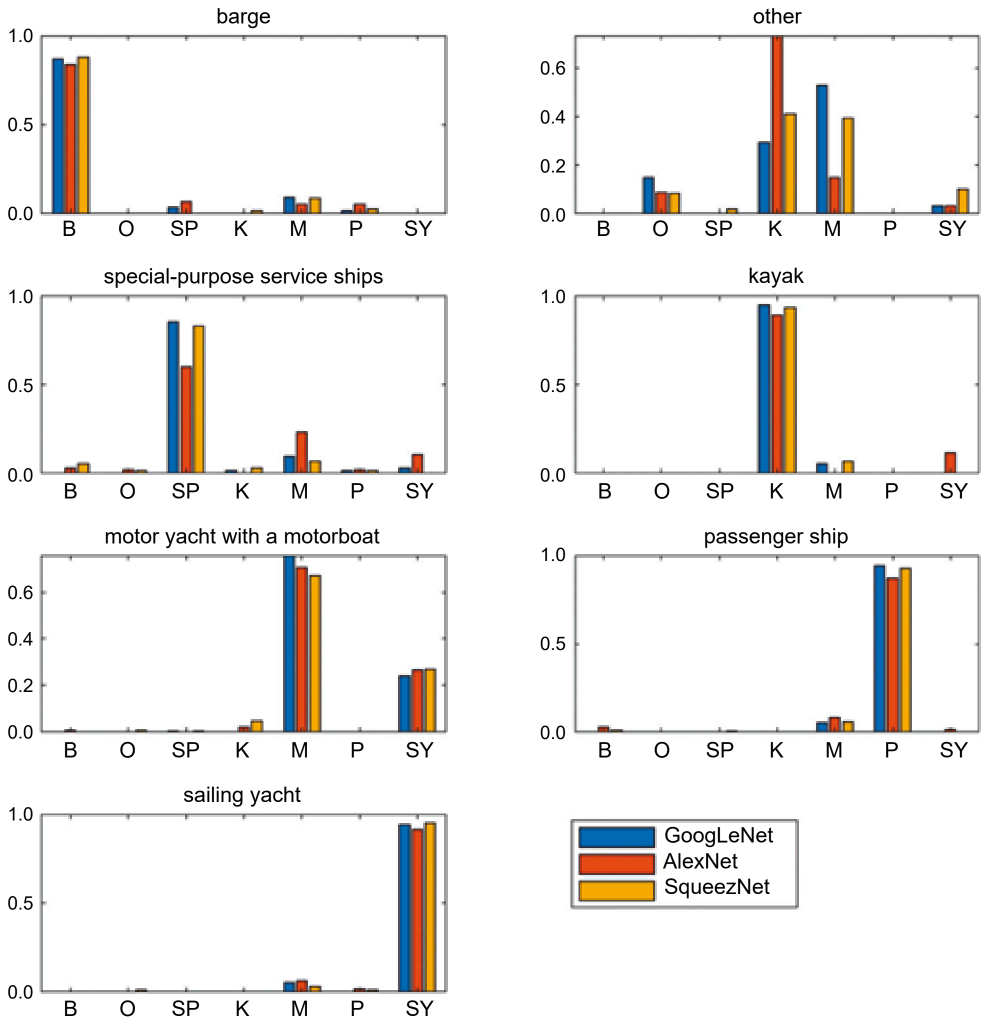

Figure 4. The correctness of classifying images as a particular class (B – barge, O – other, SP – special-purpose service ships, K – kayak, M – motor yacht, with a motorboat, P – passenger ship, SY – sailing yacht)

Acknowledgments

This work was supported by the National Centre for Research and Development (NCBR) of Poland under grant no. LIDER/17/0098/L-8/16/NCBR/2017.

References

1. Alaskar, H., Alzhrani, N., Hussain, A. & Almarshed, F. (2019) The Implementation of Pretrained AlexNet on PCG Classification. In: D.-S. Huang, Z.-K. Huang, & A. Hussain (eds) *Intelligent Computing Methodologies*, ICIC 2019. Lecture Notes in Computer Science, vol 11645. Springer, Cham, pp. 784–794, doi: 10.1007/978-3-030-26766-7_71.
2. Basheer, I. & Hajmeer, M.N. (2001) Artificial Neural Networks: Fundamentals, Computing, Design, and Application. *Journal of Microbiological Methods* 43 (1), pp. 3–31, doi: 10.1016/S0167-7012(00)00201-3.
3. Bobkowska, K. & Bodus-Olkowska, I. (2020) Potential and Use of the Googlenet Ann for the Purposes of Inland Water Ships Classification. Polish Maritime Research, 27(4), 170–178. doi: org/10.2478/pomr-2020-0077.
4. Bobkowska, K. & Wawrzyniak, N. (2019) The Hough transform in the classification process of inland ships. *Scientific Journals of the Maritime University of Szczecin, Zeszyty Naukowe Akademii Morskiej w Szczecinie* 58 (130), pp. 9–15, doi: 10.17402/331.
5. Danqing, L. (2017) *A Practical Guide to ReLU. Medium*. [Online] Available from: https://medium.com/@danqing/a-practical-guide-to-relu-b83ca804f1f7 [Accessed: March, 23 2021].
6. Espinosa, J.E., Velastin, S.A. & Branch, J.W. (2017) Vehicle Detection Using Alex Net and Faster R-CNN Deep Learning Models: A Comparative Study BT. In: H. Badioze Zaman, P. Robinson, A.F. Smeaton, T.K. Shih, S. Velastin, T. Terutoshi, A. Jaafar & N. Mohamad Ali (eds) *Advances in Visual Informatics*, pp. 3–15. Springer International Publishing.
7. Gallo, C. (2015) Artificial Neural Networks: tutorial, Encyclopedia of Information Science and Technology, Edition: 3rd Ed. (10 Volumes), USA, https://www.researchgate.net/publication/261392616.
8. Hassanpour, M. & Malek, H. (2020) Learning Document Image Features with SqueezeNet Convolutional Neural Network. *International Journal of Engineering* 33 (7), pp. 1201–1207, doi: 10.5829/ije.2020.33.07a.05.

9. Iandola, F.N., Moskewicz, M.W., Ashraf, K., Han, S., Dally, W.J. & Keutzer, K. (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 1 MB model size. CoRR, abs/1602.0. http://arxiv.org/abs/1602.07360.
10. Khan, R.U., Zhang, X. & Kumar, R. (2019) Analysis of ResNet and GoogleNet models for malware detection. *Journal of Computer Virology and Hacking Techniques* 15 (1), pp. 29–37, doi: 10.1007/s11416-018-0324-z.
11. Konovalov, D.A., Saleh, A., Bradley, M., Sankupellay, M., Marini, S. & Sheaves, M. (2019) *Underwater Fish Detection with Weak Multi-Domain Supervision*. International Joint Conference on Neural Networks (IJCNN), 1–8, doi: 10.1109/IJCNN.2019.8851907.
12. Lee, H.J., Ullah, I., Wan, W., Gao, Y. & Fang, Z. (2019) Real-Time Vehicle Make and Model Recognition with the Residual SqueezeNet Architecture. *Sensors* 19 (5), 982, doi: 10.3390/s19050982.
13. Polap, D. & Wlodarczyk-Sielicka, M. (2020) Classification of Non-Conventional Ships Using a Neural Bag-Of-Words Mechanism. *Sensors* 20 (6), 1608, doi: 10.3390/s20061608.
14. Salavati, P. & Mohammadi, H.M. (2018) *Obstacle Detection Using GoogleNet*. 8th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 326–332, doi: 10.1109/ICCKE.2018.8566315.
15. Sang-Geol, L., Yunsick, S., Yeon-Gyu, K. & Eui-Young, C. (2018) Variations of AlexNet and GoogLeNet to Improve Korean Character Recognition Performance. *Journal of Information Processing Systems* 14 (1), pp. 205–217, doi: 10.3745/JIPS.04.0061.
16. Ucar, F. & Korkmaz, D. (2020) COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Medical Hypotheses* 140, 109761, doi: 10.1016/j.mehy.2020.109761.
17. Wang, S.-H., Xie, S., Chen, X., Guttery, D.S., Tang, C., Sun, J. & Zhang, Y.-D. (2019) Alcoholism Identification Based on an AlexNet Transfer Learning Model. *Frontiers in Psychiatry* 10, 205, doi: 10.3389/fpsyt.2019.00205.
18. Wawrzyniak, N. & Stateczny, A. (2018) Automatic Watercraft Recognition and Identification on Water Areas Covered by Video Monitoring as Extension for Sea and River Traffic Supervision Systems. *Polish Maritime Research* 25, pp. 5–13, doi: 10.2478/pomr-2018-0016.
19. Wlodarczyk-Sielicka, M. & Polap, D. (2019) Automatic Classification Using Machine Learning for Non-Conventional Vessels on Inland Waters. *Sensors* (*Basel, Switzerland*) 19 (14), 3051, doi: 10.3390/s19143051.

Cite as: Bobkowska, K., Bodus-Olkowska, I. (2021) Selection of an artificial pre-training neural network for the classification of inland vessels based on their images. *Scientific Journals of the Maritime University of Szczecin, Zeszyty Naukowe Akademii Morskiej w Szczecinie* 67 (139).