Methodology for Armaments Identification Using a Neural Network

Vadym Slyusar
Chief of Research Group of Principal Researchers
Research Institute of Armament of Ukraine
Kyiv, Ukraine
swadim@ukr.net

Mykhailo Protsenko
Doctoral Student
Research Institute of Armament of Ukraine
Kyiv, Ukraine
mishapro074@gmail.com

Mariia Bondarchuk
Chief of Scientific Research Division
Research Institute of Armament of Ukraine
Kyiv, Ukraine
masha.bondarchuk@ukr.net

Authors propose to use artificial neural networks (ANNs) for armaments identification based on the analysis of digital images. They describe that this problem is caused by an increase in: the number of armaments samples, requirements to efficiency and the need to automate the process of armaments identification on the basis of digital images analysis. The authors also propose a methodology for armaments identification using ANN such as GoogleNet, NasNet-Large, Xception, Places365-Google, DenseNet-201 and others. They give an example of this methodology implementation using GoogleNet ANN, pre-trained on the ImageNet dataset. In order to solve the problem of increasing the efficiency of armaments identification based on the analysis of digital images, the authors carried out the modification of the initial full-coherent layer of pre-trained GoogleNet by 5 classes from 1000 to 5 neurons and additional training of the GoogleNet ANN. ANN training was carried out using a set of digital images by five classes: BMD-4 (airborne combat vehicle); BTR-82 (armored personnel carrier); LINSA (armored ambulance), T-72 (T-72 tank), T-90 (T-90 tank). The selected optimal training parameters were: speed (step) - 0.00005, number of epochs - 20, optimization algorithm - Adam; validation frequency - 10, packet (batch) size - 25. The effectiveness of the proposed model was tested on a set of 571 images, the total number of classes - 5. The authors selected accuracy and learning error as the main indicators of the the neural network effectiveness. The result is a new trained model with identification (classification) accuracy of validation (test) sampling of 90% and with the probability of identification error - 0.1 that proves the selection of the ANN architecture and training parameters. The use of the proposed methodology allows to automate the process of armaments identification based on the analysis of digital images of elements.

Keywords—artificial neural network, GoogleNet, identification, armaments, digital pictures, images.

I. Introduction

One of the trends in the development of machine vision and artificial intelligence in the military sphere is the development of technology for automatic armaments identification based on the digital images analysis. This problem is particularly relevant when identifying the armaments with similar appearance, for example, various tank modifications. Analysis of papers [1-8] showed that artificial neural networks (ANNs) are often used for solving the following problems: control of critical infrastructures by UAVs; remote sensing of the Earth; people (pedestrians) recognition. Thus, in [1] a model of pedestrian recognition based on convolutional neural networks (CNN) was developed. It shows that the recognition and detection of pedestrians is one of the problems for video surveillance systems; vehicle safety; robotics. A variety of factors which affect the pedestrian appearance (figure, clothing, height, lighting, background) makes this task difficult. The proposed model [1] does not solve the problem of automating the process of armaments identification based on the analysis of digital images. In paper [2] the problems of CNN architecture formation are considered. The method of random topology and multiscale matching for recognition of remotely sensed images is proposed. The proposed approach provides a better compromise between floating point operations and accuracy. Validation results of the model on the Vaihingen dataset confirm its effectiveness. The disadvantage of this model is high computational complexity, as well as unadaptability for armaments recognition based on digital images. In paper [3] a new classification scheme for hyperspectral image of remote sensing of the earth is proposed. Experimental results on three test data sets demonstrate a significant advantage of the proposed method over the known ones. The disadvantage of the method is the necessity of its adaptation for automated armaments identification based on the analysis of digital images. In papers [4-7] ANN models for UAV monitoring are considered. In papers [9-11] CNN-based models are considered. However, ANN models [4-11] are not suitable for solving the problem of armaments identification.

The analysis showed that the disadvantages of the known methods (techniques, models) are:

- lack of tested ANNs which solve the problems of armaments identification based on the analysis of digital images;

- computational complexity and instability of ANNs for different digital images of objects;

979-8-3503-9891-5/22/$31.00 ©2022 IEEE

- lack of practical application of the mathematical apparatus for the armaments identification on the basis of the digital images analysis.

Consequently, it is necessary to conduct research to develop a methodology for building neural networks for the armaments identification based on the analysis of digital images.

II. Main Part

To implement the methodology for armaments identification based on the analysis of digital images, the authors propose to use a pre-trained ANN. Setting up an ANN with repeated training is much faster than training a network with random weight coefficients. As a result, an ANN can be quickly trained to perform a new task faster and using fewer digital images and training epochs. In addition, this process of learning allows to rely on weight coefficients optimized on a significant dataset. This methodology uses the pre-trained on the ImageNet dataset [12] GoogleNet neural network [13] and Deep Learning Toolbox package of MATLAB R2020b mathematical modeling environment. For practical use the proposed methodology is presented step by step.

Step 1. Start MATLAB R2020b mathematical modeling environment (Fig. 1).

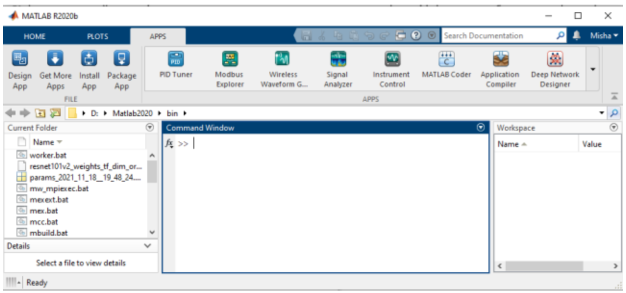

Fig. 1. MATLAB R2020b mathematical modeling environment

Step 2. Run the Deep Network Designer window in the APPS tab to download (install) the GoogleNet model for Deep Learning Toolbox neural network support package (Fig. 2).

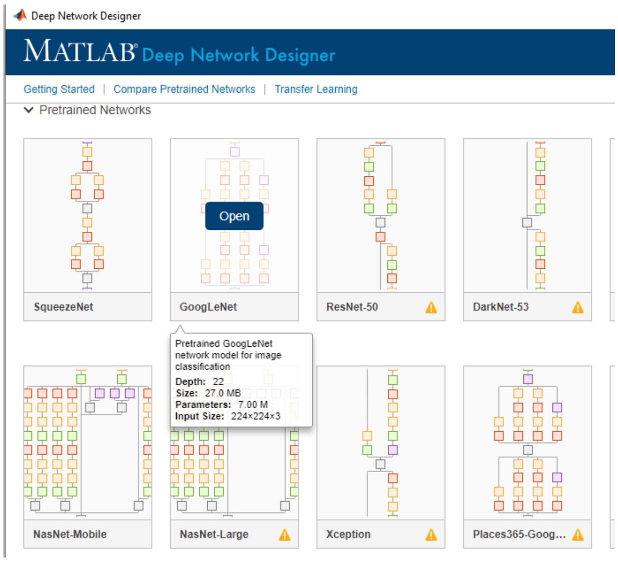

Fig. 2. MATLAB R2020b mathematical modeling environment

Step 3. Download layers of the trained ANN GoogleNet [13] (it is possible to use complex neural networks, for example: NasNet-Large [14], Xception [15], Places365-Google [16], DenseNet-201 [17], etc.). This environment allows to download pre-trained ANNs.

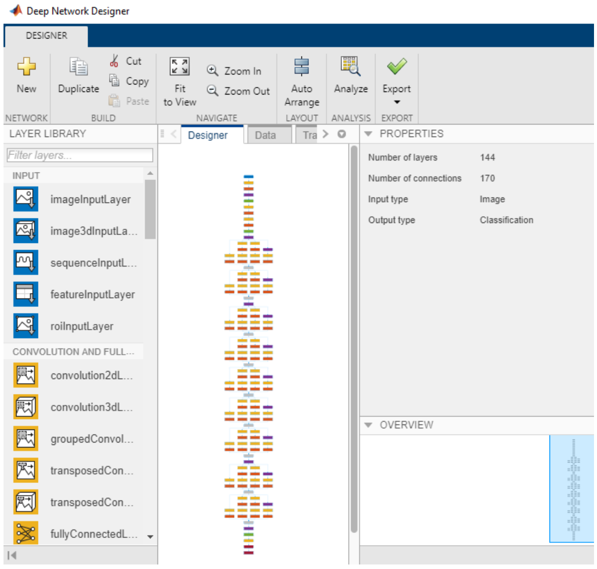

Fig. 3. Visualization of downloading the ANN model of GoogleNet using Deep Learning Toolbox neural network support package

In case of successful (full) downloading of GoogleNet model it is necessary to monitor the message (Fig. 3) with ANN properties: number of layers of ANN - 144; number of connections (links) - 170; input data type - image; output data type - classification. Since the ANN has been trained to classify images into 1000 classes on the ImageNet dataset, it must be adapted by the number of classes accordingly to the task (for example, 5 classes were taken, there can be any number of classes). In this case, the appropriate Outputsize parameter for the fc layer was set to 5.

Step 4. Use the Analyze command to interactively visualize the network architecture and detailed information about the ANN layers using the Deep Lerning Network Analyzer neural network support package (Fig. 4).

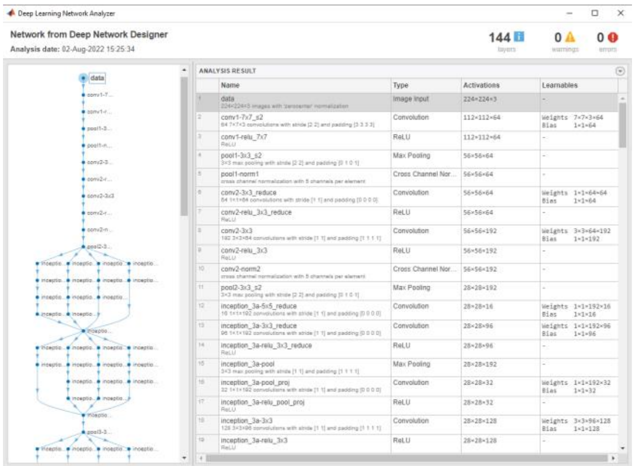

Fig. 4. Visualization of the parameters of the ANN GoogleNet layers using the Deep Lerning Network Analyzer neural network support package

The following ANN layer types are used: Image Input - input; Convolution - convolutional; Relu - relu activation functions; Cross Channel Nor... - normalization functions; Max Pooling - maxpooling functions; etc.

Step 5. Downloading new images (dataset) for training the neural network via the Data tab (Fig. 5).

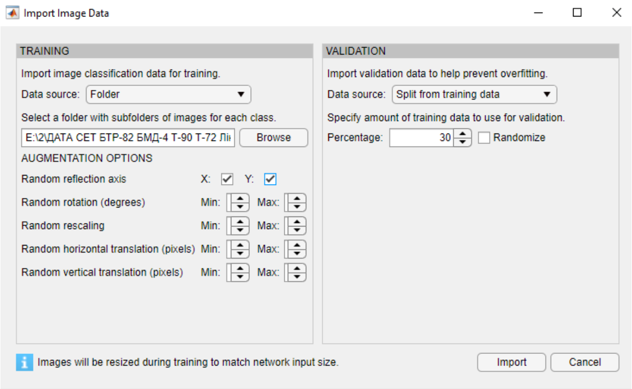

Fig. 5. Downloading new images (dataset) for NN training

At the same time, it is necessary to split the data into ANN training (Fig. 6) and ANN testing (validation) sets (Fig. 7), to analyze the existing imbalance of classes and make a decision about the expediency of data augmentation. As can be seen from Fig. 6, in this case the dataset used for ANN post-training has no significant imbalance in the number of images of different classes.

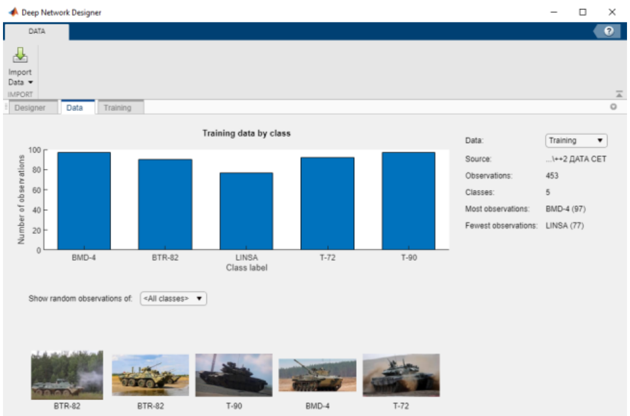

Fig. 6. Splitting dataset images into classes for ANN training

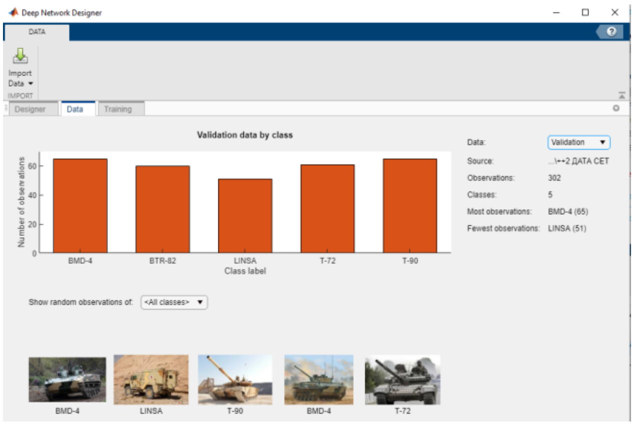

Fig. 7. Splitting dataset images into classes for ANN validation

Step 6. Selection of training parameters (Fig. 8) such as: speed (step) – 0.00005, number of epochs - 20, optimization algorithm - Adam; validation frequency - 10, package size (batch) - 25.

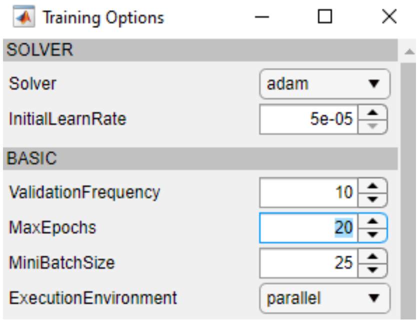

Fig. 8. ANN training parameters

Increasing the number of images helps to prevent the network from overtraining. The parallel method was used to train the ANN simultaneously on the CPU and the graphics processing unit (GPU). The ANN training process and its characteristics are shown in Fig. 9.

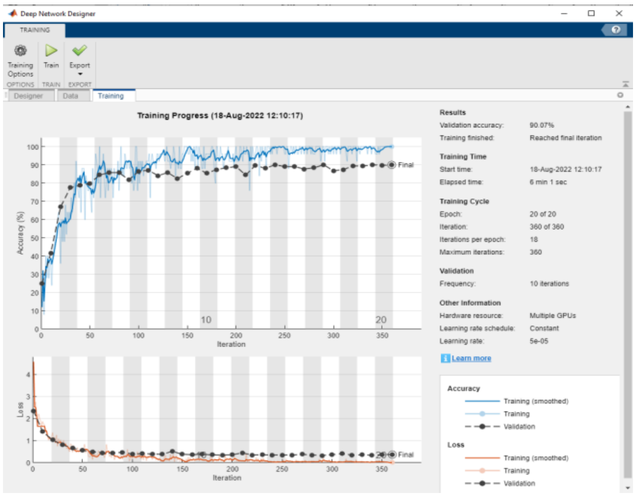

Fig. 9. Assessment of classification accuracy of validation (test) sampling - 90%, Epoch 20, Iteration 360

The result of the post-trained ANN on the recognition of classes of images is shown in Fig. 10.

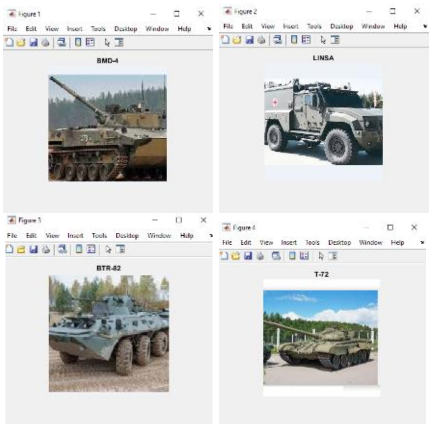

Fig. 10. Results of armaments identification on the basis of ANN digital images analysis

In Fig. 10 you can see the results of application of the ANN for armaments identification on the basis of analysis of digital images: BMD-4 (airborne combat vehicle); LINSA (armored ambulance), BTR-82 (armored personnel carrier); T-72 (T-72 tank). The ANN recognized all images correctly. The accuracy of this ANN at the 20th epoch of training reached 90%. The proposed GoogleNet model was compared with the AlexNET model. The ANN training process and its characteristics are shown in Fig. 11.

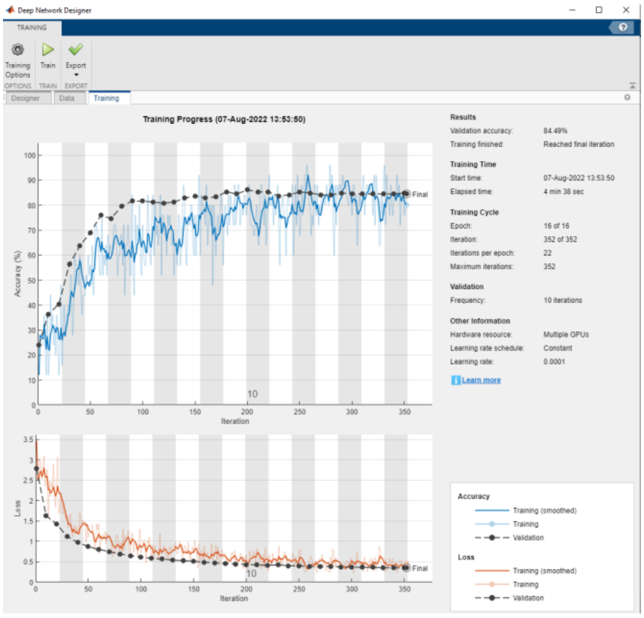

Fig. 11. Assessment of classification accuracy of validation (test) sampling – 84%, Epoch 18, Iteration 352

Classification accuracy of the validation (test) sampling was 84 %. Accordingly, the best results were shown by the GoogleNet model.

III. Conclusion

1. In order to solve the problem of armaments identification based on the analysis of digital images, the methodology for armaments identification using a neural network was developed and tested on a specific example. The resulting model allows to automate the process of armaments identification based on the analysis of digital images.

2. Analysis of possible implementations showed that GoogleNet ANN trained by ImageNet set of images is suitable for this task. This model was used as a basic one. To increase the efficiency we cut the initial full-coherent layer from 1000 to 5 neurons and additionally trained the resulting model by a set of images of five classes: BMD-4 (airborne combat vehicle); BTR-82 (armored personnel carrier); LINSA (armored ambulance), T-72 (T-72 tank), T-90 (T-90 tank). By selection of optimal parameters (speed (step) - 0,00005, number of epochs - 20, optimization algorithm - Adam; validation frequency - 10, packet (batch) size - 25) the authors received a new model with classification accuracy of validation (test) sampling of 90 %.

3. The use of this model makes it possible to automate the process of armaments identification on the basis of digital images analysis.

4. The ANN training was conducted on digital images of high contrast and clarity. The shooting was done in the daytime that is why high accuracy of image recognition was obtained. For other types of images (shooting conditions) recognition accuracy by classes may vary, which requires additional research.

5. The limitations of the proposed model are that it is adapted to recognize objects in digital images only by five classes. It did not take into account the orientation of objects in the images and the translational invariance of the ANN.

6. For further development of the proposed model the authors plan to:

- evaluate the accuracy of the developed model for different object recognition conditions and different ANN models like NasNet-Large [14], Xception [15], Places365-Google [16], DenseNet-201 [17], EfficientNet [18];

- increase the accuracy and performance of the ANN and reduce the volume of calculations;

- increase the number of classes for classification objects and the volume of dataset for ANN post-training;

- study the proposed architecture with other activation and maxpooling functions (among those proposed in [19]).

References

[1] P. Sermanet, K. Kavukcuoglu, S. Chintala, , Y. Lecun, “Pedestrian Detection with Unsupervised Multi-stage Feature Learning,” 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3626–3633. https://doi.org/10.1109/CVPR.2013.465.

[2] J. Li, M. Weinmann, X. Sun, W. Diao, Y. Feng, K. Fu, “Random Topology and Random Multiscale Mapping: An Automated Design of Multiscale and Lightweight Neural Network for Remote-Sensing Image Recognition,” in IEEE Transactions on Geoscience and Remote Sensing, 2021, pp. 1–17. https://doi.org/10.1109/TGRS.2021.3102988.

[3] S. Yang, J. Hou, Y. Jia, S. Mei, Q. Du, “Superpixel-Guided Discriminative Low-Rank Representation of Hyperspectral Images for Classification,” in IEEE Transactions on Image Processing, vol. 30, 2021, pp. 8823–8835. https://doi.org/10.1109/TIP.2021.3120675.

[4] V. Slyusar, M. Protsenko, A. Chernukha, S. Gornostal, S.Rudakov, S. Shevchenko, O. Chernikov, N. Kolpachenko, V. Timofeyev, R. Artiukh, “Construction of an advanced method for recognizing monitored objects by a convolutional neural network using a discrete wavelet transform,” Eastern-European Journal of Enterprise Technologies, 4(9(112), 2021, pp. 65–77. https://doi.org/10.15587/1729-4061.2021.238601.

[5] V. Slyusar, M. Protsenko, A. Chernukha, P. Kovalov, P. Borodych, S. Shevchenko, O. Chernikov, S. Vazhynskyi, O. Bogatov, K. Khrustalev, “Improvement of the model of object recognition in aero photographs using deep convolutional neural networks.,” Eastern-European Journal of Enterprise Technologies, 5(2(113), 2021, pp. 6–21. https://doi.org/10.15587/1729-4061.2021.243094.

[6] V. Slyusar, M. Protsenko, A. Chernukha, V. Melkin, O. Petrova, M. Kravtsov, S. Velma, N. Kosenko, O. Sydorenko, M. Sobol, “Improving a neural network model for semantic segmentation of images of monitored objects in aerial photographs,” Eastern-European Journal of Enterprise Technologies, 6(2(114)), 2021, pp. 86–95. https://doi.org/10.15587/1729-4061.2021.248390.

[7] V. Slyusar, M. Protsenko, A. Chernukha, V. Melkin, O. Biloborodov, M. Samoilenko, O. Kravchenko, H. Kalynychenko, A. Rohovyi, M. Soloshchuk, “Improving the model of object detection on aerial photographs and video in unmanned aerial systems,” Eastern-European Journal of Enterprise Technologies,

1(9(115)), 2022. pp. 24–34. https://doi.org/10.15587/1729-4061.2022.252876.

[8] D. Lian, “Multiview Multitask Gaze Estimation With Deep Convolutional Neural Networks,” in IEEE Transactions on Neural Networks and Learning Systems, 10/30, 2019, pp. 3010–3023. https://doi.org/10.1109/TNNLS.2018.2865525.

[9] H. Li, , J. Li, X. Han, “Robot Vision Model Based on Multi-Neural Network Fusion,” 2019 IEEE 3rd Information Technology Networking Electronic and Automation Control Conference (ITNEC), 2019, pp. 2571–2577. https://doi.org/10.1109/ITNEC.2019.8729210.

[10] Z. Zhong-Qiu, Z. Peng, X. Shou-tao, W. Xindong, “Object Detection with Deep Learning,” A Review (cs.CV), 2019, pp. 1–21. https://arxiv.org/pdf/1807.05511.pdf

[11] J. Kim, J. Koh, B. Lee, S. Yang, J. Choi, “Video Object Detection Using Object's Motion Context and Spatio-Temporal Feature Aggregation,” 2020 25th International Conference on Pattern Recognition (ICPR), 2020, pp. 1604–1610. doi: https://doi.org/10.1109/ICPR48806.2021.9412715.

[12] ImageNet official site [Electronic resource]. – Access mode: https://www.image-net.org/.

[13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, “Goint Deep with Convolutions,” Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1–12. https://arxiv.org/abs/1409.4842.

[14] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8697–8710, https://arxiv.org/abs/1707.07012.

[15] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1–8. https://arxiv.org/pdf/1610.02357.pdf.

[16] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, A. Torralba, “Places: A 10 million Image Database for Scene Recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017, pp. 1–14. http://places2.csail.mit.edu/PAMI_places.pdf

[17] H. Gao, L. Zhuang, Q. Kilian, “Weinberger, Laurens van der Maaten. Densely Connected Convolutional Networks,” Computer Vision and Pattern Recognition (CVPR), 2018, pp. 1–9. https://arxiv.org/pdf/1608.06993.pdf

[18] M. Tan, Q.Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning Research 97: 2019, pp. 6105-6114 Available from https://proceedings.mlr.press/v97/tan19a.html.

[19] V. Slyusar, “Architectural and mathematical foundations of improving neural networks for image classification.” Artificial intelligence, 2022, pp. 127–138. https://doi.org/10.15407/jai2022.01.127.

979-8-3503-9891-5/22/$31.00 ©2022 IEEE