

КЛАСИФІКАЦІЯ ТЕКСТІВ ЯК ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ НЕЙРОМЕРЕЖ, ПОПЕРЕДНЬО НАВЧЕНИХ НА ДАТАСЕТІ ІМАГЕНЕТ

Слюсар Вадим Іванович

доктор технических наук , профессор



Bag of Words

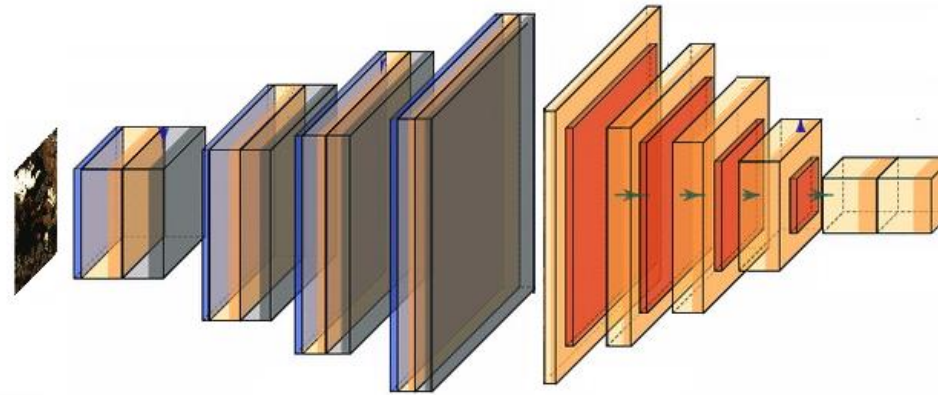
Базова згорткова неймережа класифікації текстів як зображень

Середньокласова точність класифікації за метрикою **Balanced Recall 88.6 %**.

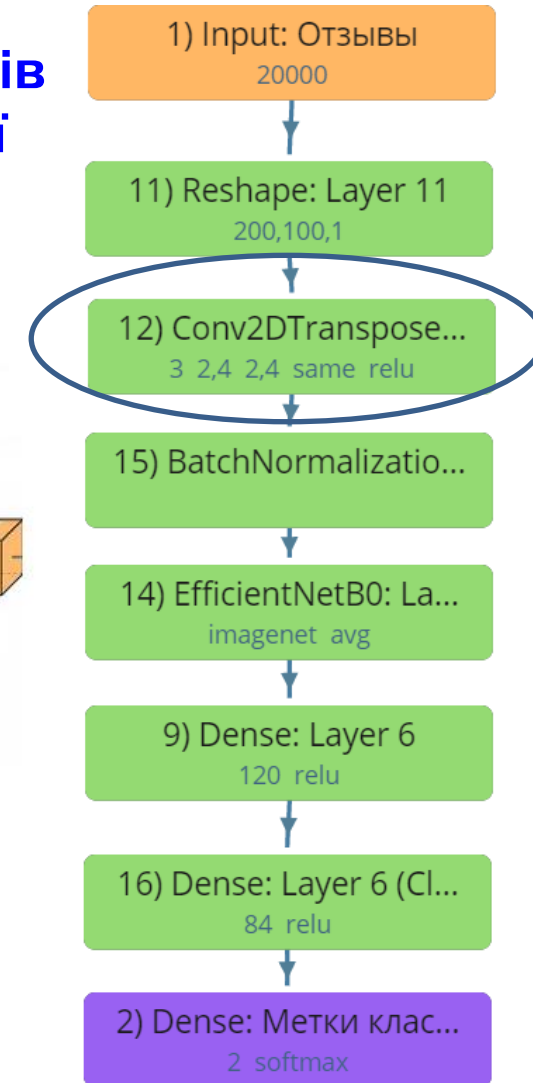
Досягти точності **88.7 %** дозволив перехід у згортковому шарі Conv2D до 64 фільтрів у поєднанні з ваговим ядром формату 2x2. Крім того, було збільшено кількість нейронів у першому повнозв'язному шарі Dense до 256.

Неймережа класифікації текстів на основі попередньо навченої EfficientNetB0

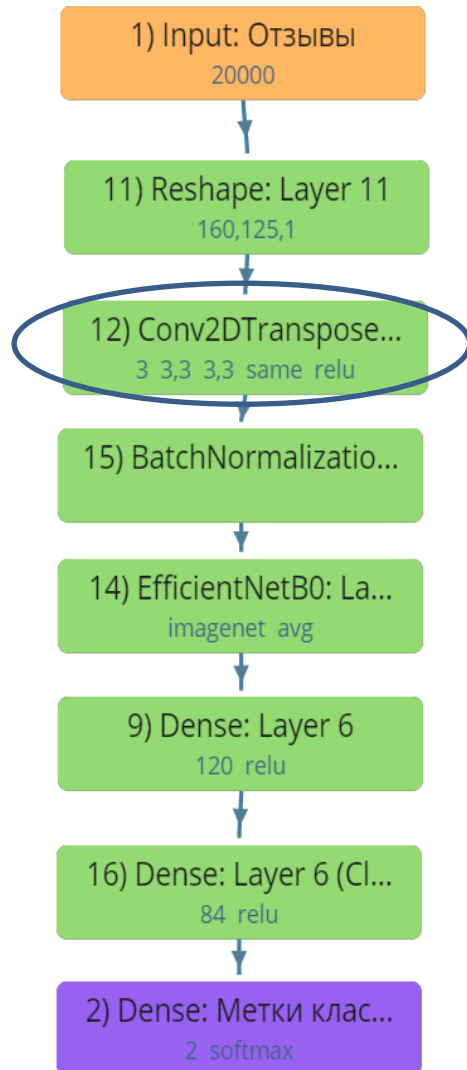
91.7 %.



94.5 %.



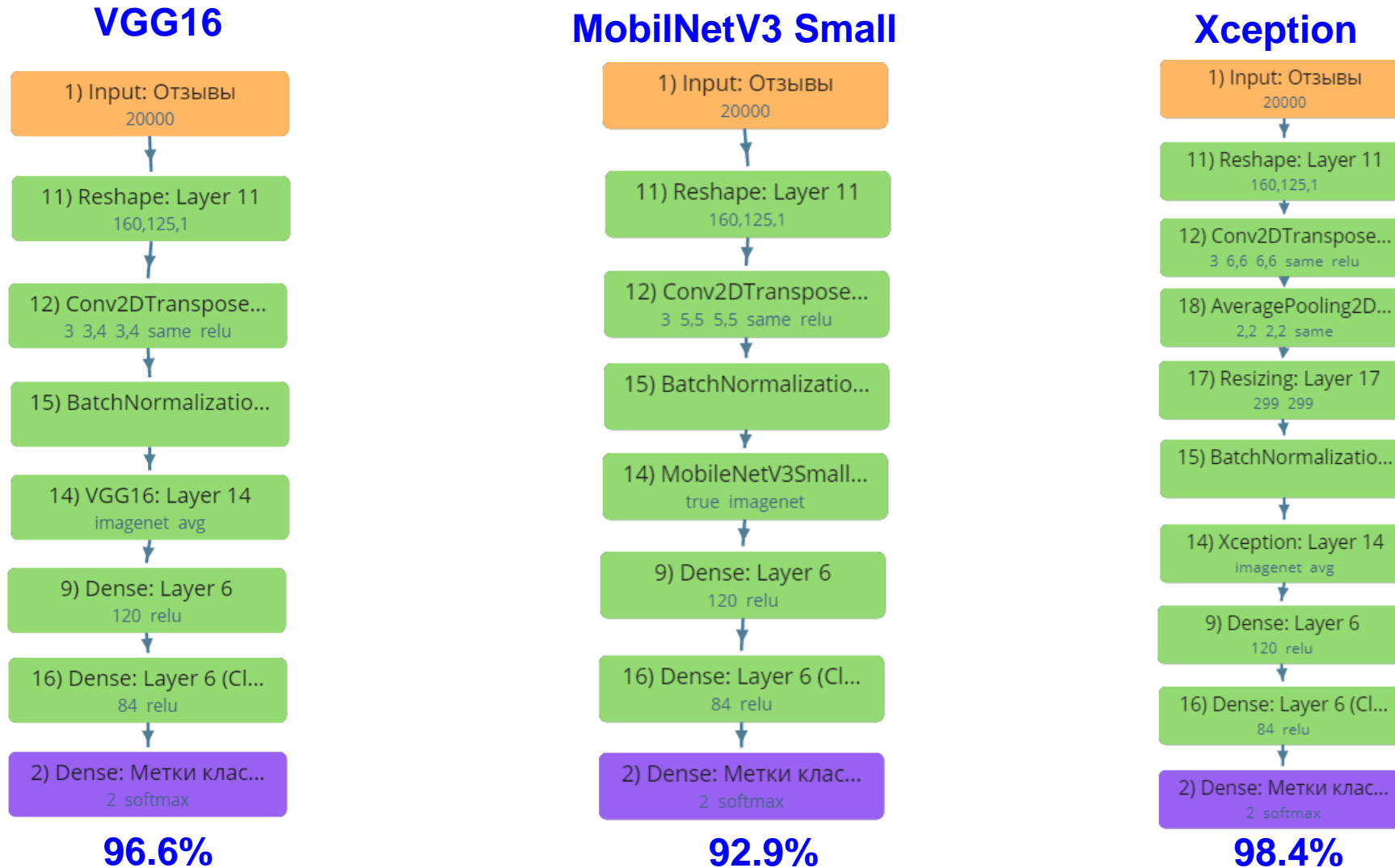
EfficientNetB0



99.5 % !!!



Досліджені варіанти неймереж



Метод підготовки датасета Word2vec

Xception



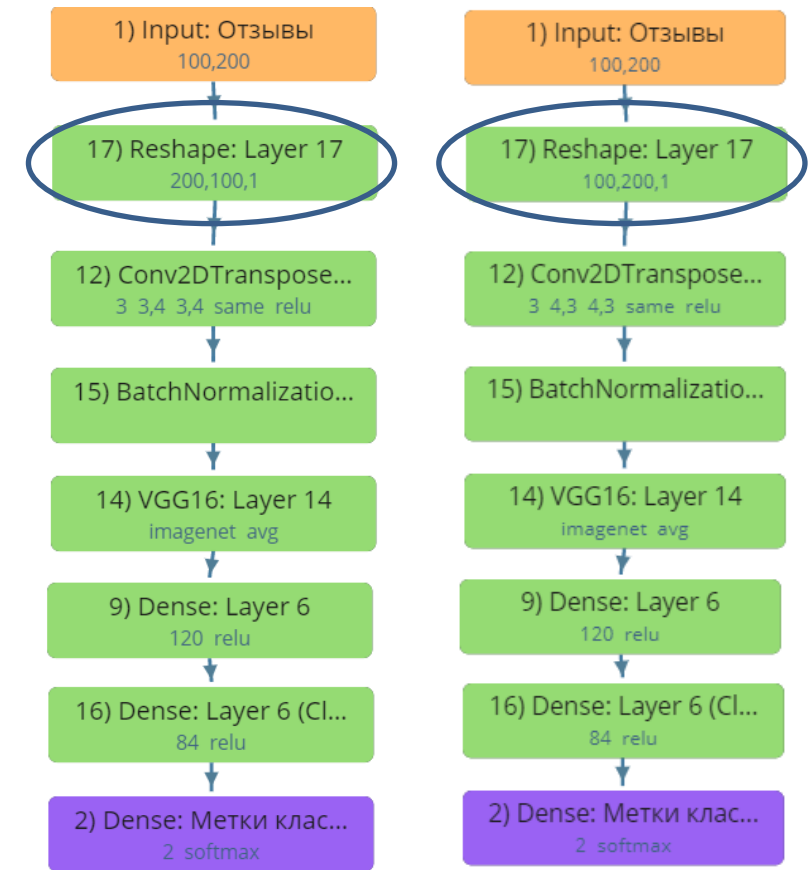
79 %

EfficientNetB0



69 %

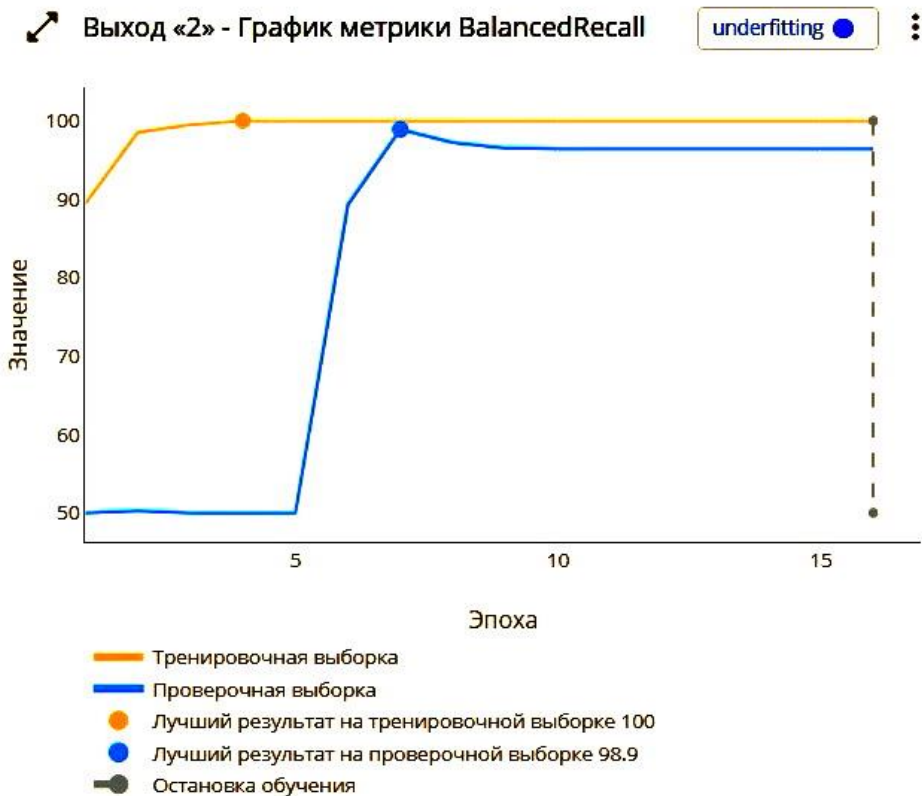
VGG16



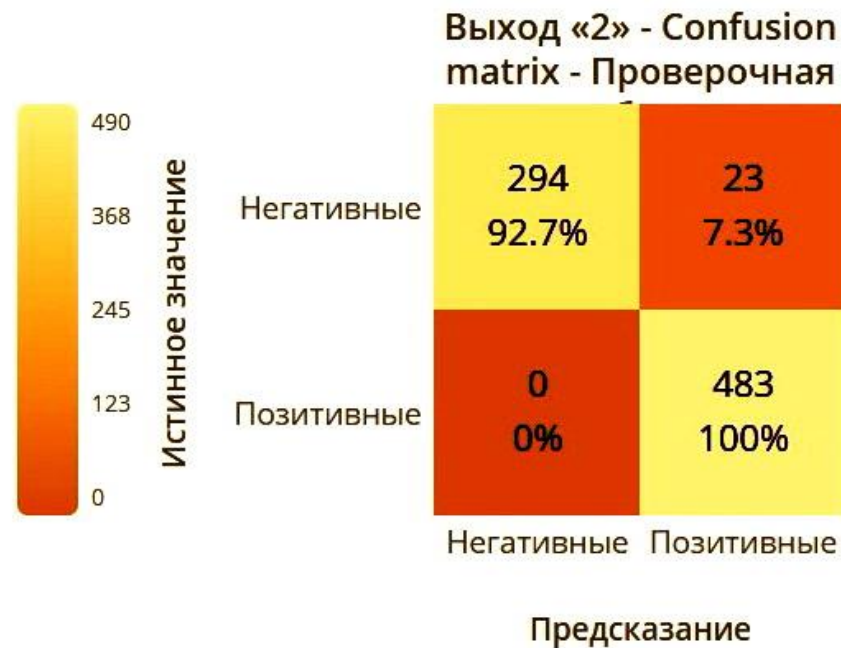
71.9 %

Метод підготовки датасета Embedding

Хсепшн
Подвійний Embedding



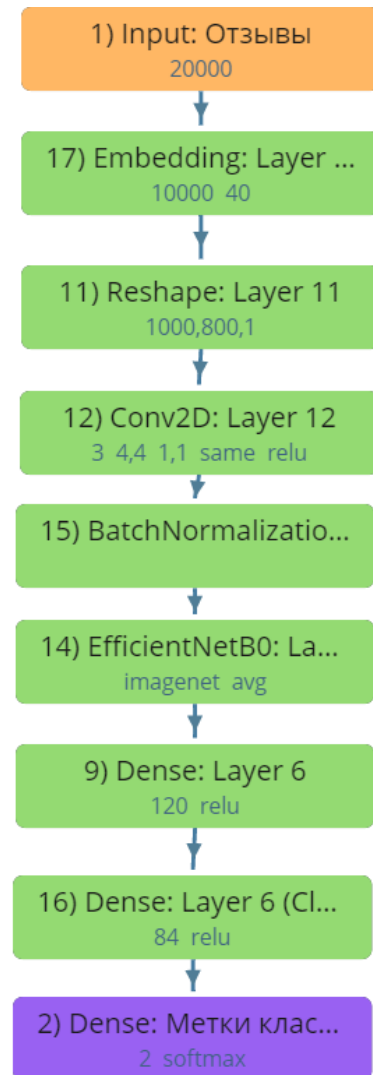
Точність 98.9% на тестовій вибірці



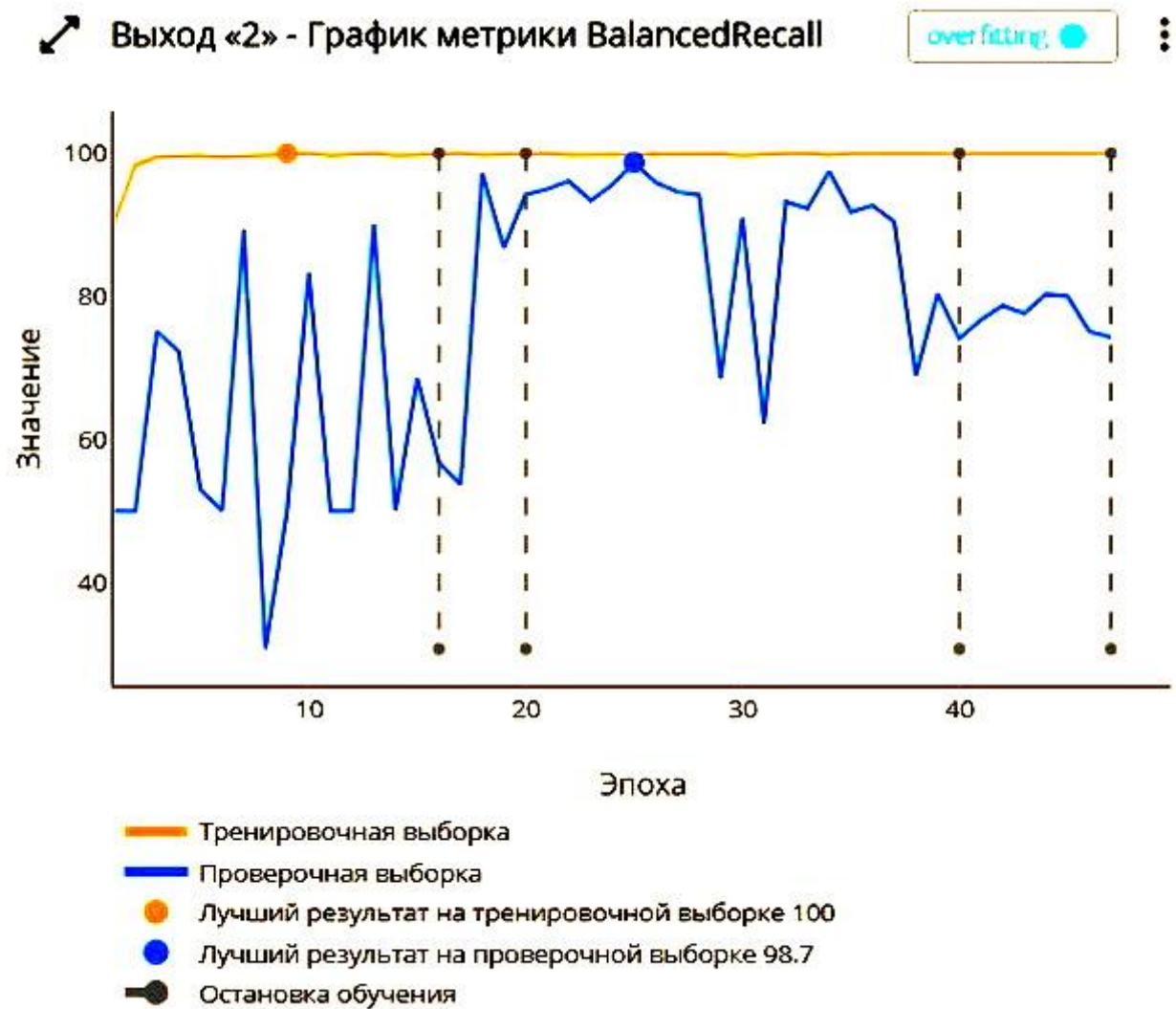
Матрица невідповідності

Комбінація Bag of Words та Embedding

EfficientNetB0



98.7%.



Висновки

1. Проведені дослідження засвідчили можливість ефективної реалізації завдань класифікації текстів за допомогою попередньо навчених згорткових нейронних мереж, що застосовуються для обробки зображень
2. Серед проаналізованих варіантів попередньої підготовки текстового датасету найкращі результати з точності класифікації текстів нейромережами з обробки зображень в середньому забезпечує метод мішка слів (BoW). При використанні попередньо навченої на датасеті ImageNet нейромережі EfficientNetB0 такий підхід дозволив отримати середньокласову точність класифікації відгуків **99.5 %**.
3. Процедура ембеддінгу дещо поступається методу BoW.
4. Найменш успішним варіантом цифровізації текстів виявився метод Word2vec.
5. Заслуговує на увагу запропонований підхід щодо комбінації методу підготовки текстового датасету BoW з додатковою процедурою Embedding у складі нейромережі. Таке поєднання у випадку EfficientNetB0 дозволило досягти порівняно високої точності 98.7 %.

References

1. Benarab Charaf Eddine. Classifying Textual Data with pretrained Vision Models through Transfer Learning and Data Transformations. // Feb. 7, 2022, 7 p. arXiv:2106.12479v4. <https://arxiv.org/pdf/2106.12479.pdf>.
2. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248 – 255. Ieee, 2009.
3. M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proc. of International Conference on Machine Learning*, 2019, pp. 6105-6114.
4. Sandler, M., Howard, A., Zhu, M., et al. (2018) Mobilenetv2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. DOI: 10.1109/CVPR.2018.00474.
5. Howard, A., Sandler, M., Chu, G., et al. (2019) Searching for Mobilenetv3. *Proceedings of the IEEE International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 1314-1324. DOI: 10.1109/ICCV.2019.00140.
6. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800-1807, DOI: 10.1109/CVPR.2017.195.

References (cont.)

7. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016 pp. 2818-2826. DOI: 10.1109/CVPR.2016.308.
8. H. Qassim, A. Verma and D. Feinzimer (2018), Compressed residual-VGG16 CNN model for big data places image recognition, *Computing and Communication Workshop and Conference (CCWC) 2018 IEEE 8th Annual*, 169-175.
9. Vadym Slyusar, Mykhailo Protsenko, Anton Chernukha, Vasyl Melkin, Olena Petrova, Mikhail Kravtsov, Svitlana Velma, Nataliia Kosenko, Olga Sydorenko, Maksym 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.DOI: 10.15587/1729-4061.2021.248390.
10. Слюсар В.І. Архітектурно-математичні основи удосконалення нейронних мереж з класифікації зображень. // *Штучний інтелект*, 2022, №1.- С. 127 - 138. DOI: 10.15407/jai2022.01.127.
11. Слюсар В.И., Слюсарь И.И. Львы зоопарка нейростей.// *Нейромережні технології та їх застосування НМТіЗ-2021: збірник наукових праць XX Міжнародної наукової конференції «Нейромережні технології та їх застосування НМТіЗ-2021».* - Краматорськ: Донбаська державна машинобудівна академія. – 8 - 9 грудня 2021. - С. 129 -133. DOI: 10.13140/RG.2.2.17187.58405

THANK YOU

FOR YOUR ATTENTION !

swadim@ukr.net

www.slysuar.kiev.ua

<https://scholar.google.com.ua/citations?hl=ru&user=wSegaWsAAAAJ>

<https://orcid.org/0000-0002-2912-3149>

<https://www.scopus.com/authid/detail.uri?authorId=7004240035>

<https://www.researchgate.net/profile/Vadym-Slyusar>