

Вестник НГУЭУ. 2022. № 2. С. 228–238

Vestnik NSUEM. 2022. No. 2. P. 228–238

Научная статья

УДК 004.89

DOI: 10.34020/2073-6495-2022-2-228-238

**ЭФФЕКТИВНОСТЬ МОДЕЛЕЙ TENSORFLOW
В ПРИЛОЖЕНИИ К ЗАДАЧЕ ОБНАРУЖЕНИЯ ГЛАЗ
НА ФОТОГРАФИИ**

Бурдуковский Станислав Олегович

Новосибирский государственный университет экономики и управления «НИНХ»

burdukowsky_stas@mail.ru

Аннотация. В данной статье автор проводит сравнение эффективности моделей обнаружения TensorFlow при решении задачи детектирования области глаз на фотографии лица человека. Для этого был проведен ряд экспериментов двух видов: дообучение предварительно обученной модели обнаружения и обучение модели с нуля. Для формирования обучающей и оценочных выборок использовались изображения лиц из открытой базы данных Flickr-Faces-HQ Dataset. В статье приведены параметры обучения, показаны графики потерь классификации и локализации, произведена оценка точности работы обученных моделей, а также продемонстрирована работа обученной с нуля модели обнаружения «SSD MobileNet V2 FPNLite 320×320», которая получила самые высокие оценки точности после обучения с нуля и после дообучения. Для программ с требованием по IoU обнаруживаемых объектов больше 0,5 точность работы модели составила 99,9 %. Результаты экспериментов могут быть использованы в различных исследованиях, в которых применяется платформа TensorFlow для обнаружения объектов на изображениях, и производится обнаружение объектов только одного класса.

Ключевые слова: обнаружение объектов, машинное обучение, глаза, искусственная нейронная сеть, TensorFlow, потери классификации, потери локализации, точность обнаружения

Для цитирования: Бурдуковский С.О. Эффективность моделей TensorFlow в приложении к задаче обнаружения глаз на фотографии // Вестник НГУЭУ. 2022. № 2. С. 228–238. DOI: 10.34020/2073-6495-2022-2-228-238.

Original article

THE EFFICIENCY OF THE TENSORFLOW MODELS IN THE APPLICATION TO THE TASK OF DETECTION OF EYES IN THE PHOTO

Burdukowsky Stanislav O.

Novosibirsk State University of Economics and Management

burdukowsky_stas@mail.ru

Abstract. In this article, the author compares the effectiveness of TensorFlow detection models in solving the problem of detecting areas with eyes in human face photo. Experiments of two types were carried out: additional training of a pretrained detection model and training of the model from scratch. Face images from Flickr-Faces-HQ Dataset were used to form training and evaluation samples. The article describes the training parameters, shows classification and localization loss graphs, assesses the accuracy of the trained models, and also demonstrates the operation of the “SSD MobileNet V2 FPNLite 320×320” detection model trained from scratch, which received the highest accuracy scores after additional training and training from scratch. For programs with a requirement for IoU of detected objects greater than 0.5, the accuracy of the model is 99.9 %. The results of the experiments can be applied in various researches, that use the TensorFlow platform to detect objects in images, and only one class of objects is detected.

Keywords: object detection, machine learning, eyes, artificial neural network, TensorFlow, classification loss, localization loss, detection accuracy

For citation: Burdukowsky S.O. The efficiency of the TensorFlow models in the application to the task of detection of eyes in the photo. *Vestnik NSUEM*. 2022; (2): 228–238. (In Russ.). DOI: 10.34020/ 2073-6495-2022-2-228-238.

Введение

В данной статье автор проводит сравнение эффективности моделей обнаружения TensorFlow при решении задачи детектирования области глаз на фотографии лица человека.

TensorFlow – это комплексная платформа с открытым исходным кодом для машинного обучения, разработанная компанией Google [21]. TensorFlow предоставляет комплекс моделей обнаружения, предварительно обученных на наборе данных COCO 2017 [14]. Пользователи могут использовать эти модели для решения своих задач в готовом виде, провести дообучение для распознавания классов объектов, отсутствующих в наборе данных COCO 2017, или же обучить эти модели с нуля.

Представленные в статье результаты являются частью научного исследования, посвященного разработке информационной системы автоматического диагностирования косоглазия. Первый этап при обработке входного изображения – выделение на нем областей, на которых присутствуют глаза. Сравнение эффективности моделей обнаружения TensorFlow было произведено с целью определения наиболее подходящей для решения задачи

детектирования области глаз. Для этого проводились эксперименты двух видов:

- дообучение предварительно обученной модели обнаружения для распознавания области глаз,
- обучение модели для распознавания области глаз с нуля.

Эти эксперименты были проведены для каждой модели обнаружения, предоставляемой платформой TensorFlow, а именно:

- CenterNet HourGlass104 512×512,
- CenterNet Resnet50 V1 FPN 512×512,
- CenterNet Resnet101 V1 FPN 512×512,
- CenterNet Resnet50 V2 512×512,
- CenterNet MobileNetV2 FPN 512×512,
- EfficientDet D0 512×512,
- SSD MobileNet v2 320×320,
- SSD MobileNet V1 FPN 640×640,
- SSD MobileNet V2 FPNLite 320×320,
- SSD ResNet50 V1 FPN 640×640 (RetinaNet50),
- SSD ResNet101 V1 FPN 640×640 (RetinaNet101),
- SSD ResNet152 V1 FPN 640×640 (RetinaNet152),
- Faster R-CNN ResNet50 V1 640×640,
- Faster R-CNN ResNet101 V1 640×640,
- Faster R-CNN ResNet152 V1 640×640,
- Faster R-CNN Inception ResNet V2 640×640.

Платформа TensorFlow была выбрана по результатам анализа различных инструментов глубокого обучения (Chainer, Keras, Neural Network Toolbox, Microsoft Cognitive Toolkit, Apache MXNet, PyTorch, TensorFlow и Mathematica). Преимущества платформы TensorFlow: хорошо поддерживаемая, открытая и поддерживающая большое количество технологий, таких как CUDA (Compute Unified Device Architecture), автоматическое дифференцирование, предтренированные модели, рекуррентные нейронные сети, сверточные нейронные сети, машина Больцмана, глубокая сеть доверия, вычисления на компьютерном кластере [1].

Набор данных для обучения

Для обучения модели были взяты изображения лиц из базы данных Flickr-Faces-HQ Dataset. Эта БД содержит 70 000 фотографий лиц, которые имеют отличия с точки зрения возраста человека, этнической принадлежности и фона изображения. Также имеется большой охват аксессуаров, таких как головные уборы, украшения и прочее [15].

Для проведения экспериментов была использована выборка объемом 25 000 изображений. Из них 2500 составили оценочную выборку (те изображения, которые использовались для оценки обученной модели и не участвовали в обучении), а 22 500 – обучающую. Это соотношение составляет 1 : 9, и было выбрано на основании опыта специалистов в области науки о данных. Специалисты, основываясь на результатах экспериментов, указывают на то, что соотношение 1 : 9 является наиболее эффективным [19, 20].

На основе всех 25 000 изображений и разметки (координат объектов на изображениях) были сгенерированы с помощью скрипта на языке программирования Python файлы аннотаций для обучающей и оценочной выборок в формате CSV. Файл аннотации – это файл, содержащий информацию (координаты прямоугольных областей) об объектах (в данном случае о глазах) на конкретном изображении. Далее данные для обучения и оценки были сконвертированы в формат TFRecord с помощью инструментов, предоставляемых платформой TensorFlow.

Обучение

Количество шагов обучения – 50 000. Каждый шаг обучения представляет собой обучение на восьми изображениях (размер партии данных). Это значит, что обучение было произведено на 400 000 экземплярах изображений. Так как всего изображений в обучающей выборке 22 500, то количество эпох обучения составило приблизительно 17,8.

Размер партии данных подобран эмпирическим путем исходя из технических данных компьютера, на котором проводились эксперименты. Размер партии влияет только на производительность обучения и не влияет на результаты. Количество шагов обучения (50 000) было определено также эмпирически, на основе нескольких экспериментов. В каждом эксперименте примерно на сорокатысячном шаге потери классификации и локализации стабилизировались и значительных изменений в сторону уменьшения потерь не наблюдалось (рис. 1, 2).

В процессе обучения вычислялись и записывались суммы значений потерь классификации и локализации. Потери классификации – это мера того, как сильно модель ошибается при распознавании глаза и фона, а потери локализации – это мера того, как сильно модель ошибается при локализации глаза. Для примера представлен журнал обучения с нуля модели «SSD MobileNet V2 FPNLite 320×320» (последняя тысяча шагов). Сумма потерь указана в поле «loss» (см. рис. 1).

```
I0415 08:11:27.149714 140355650852672 model_lib_v2.py:682] Step 49000 per-step time 0.076s loss=0.350
INFO:tensorflow:Step 49100 per-step time 0.079s loss=0.218
I0415 08:11:35.791534 140355650852672 model_lib_v2.py:682] Step 49100 per-step time 0.079s loss=0.218
INFO:tensorflow:Step 49200 per-step time 0.137s loss=0.286
I0415 08:11:44.242342 140355650852672 model_lib_v2.py:682] Step 49200 per-step time 0.137s loss=0.286
INFO:tensorflow:Step 49300 per-step time 0.080s loss=0.287
I0415 08:11:52.505284 140355650852672 model_lib_v2.py:682] Step 49300 per-step time 0.080s loss=0.287
INFO:tensorflow:Step 49400 per-step time 0.079s loss=0.179
I0415 08:12:00.891387 140355650852672 model_lib_v2.py:682] Step 49400 per-step time 0.079s loss=0.179
INFO:tensorflow:Step 49500 per-step time 0.078s loss=0.248
I0415 08:12:09.139921 140355650852672 model_lib_v2.py:682] Step 49500 per-step time 0.078s loss=0.248
INFO:tensorflow:Step 49600 per-step time 0.094s loss=0.190
I0415 08:12:17.420404 140355650852672 model_lib_v2.py:682] Step 49600 per-step time 0.094s loss=0.190
INFO:tensorflow:Step 49700 per-step time 0.092s loss=0.207
I0415 08:12:25.767324 140355650852672 model_lib_v2.py:682] Step 49700 per-step time 0.092s loss=0.207
INFO:tensorflow:Step 49800 per-step time 0.077s loss=0.242
I0415 08:12:34.045423 140355650852672 model_lib_v2.py:682] Step 49800 per-step time 0.077s loss=0.242
INFO:tensorflow:Step 49900 per-step time 0.080s loss=0.223
I0415 08:12:42.213774 140355650852672 model_lib_v2.py:682] Step 49900 per-step time 0.080s loss=0.223
INFO:tensorflow:Step 50000 per-step time 0.077s loss=0.205
I0415 08:12:50.327841 140355650852672 model_lib_v2.py:682] Step 50000 per-step time 0.077s loss=0.205
```

Рис. 1. Журнал обучения с нуля модели «SSD MobileNet V2 FPNLite 320×320»
Journal of learning from scratch the model “SSD MobileNet V2 FPNLite 320×320”

Также в процессе обучения платформа TensorFlow позволяет просматривать по отдельности графики потерь классификации и локализации. Далее представлены такие графики для обучения с нуля модели «SSD MobileNet V2 FPNLite 320×320» (рис. 2).

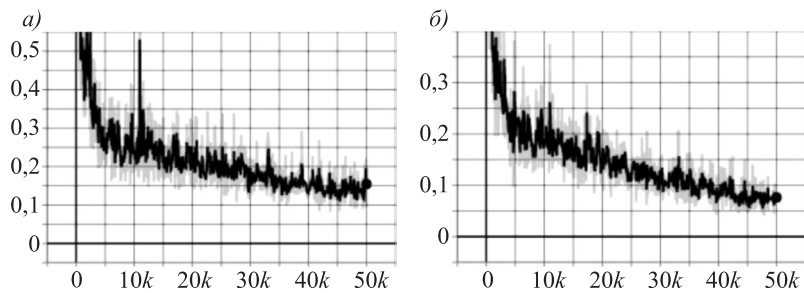


Рис. 2. Пример графиков потерь классификации (а) и локализации (б)
An example of classification and localization loss plots

Ось абсцисс показывает номер шага обучения, а ось ординат – величину потери классификации или локализации. Буква «k» на графиках обозначает «тысяч».

Функция потерь измеряет разницу между выходным и целевым значением модели (выходным являются координаты объекта на изображении). На основании величины потерь платформа TensorFlow во время обучения корректирует параметры модели. Скачки на графиках объясняются тем, что на каждом шаге осуществляется коррекция параметров модели, и она не всегда улучшает точность работы. Но в целом, как видно на рис. 2, динамика изменения потерь отрицательная.

В качестве метрики для модели обнаружения объектов используется значение IoU (Intersection over Union). Оно обозначает отношение площади пересечения оригинальной рамки объекта на изображении и рамки объекта, найденной моделью, к площади области, полученной путем объединения этих рамок:

$$IoU = \frac{S_{\text{пер}}}{S_{\text{об}}}.$$

На рис. 3 представлена наглядная иллюстрация IoU.

Для оценки точности работы модели обнаружения объектов используется среднее арифметическое долей верно обнаруженных объектов с IoU больше 50 %, 55, 60, 65, 70, 75, 80, 85, 90 и 95 %:

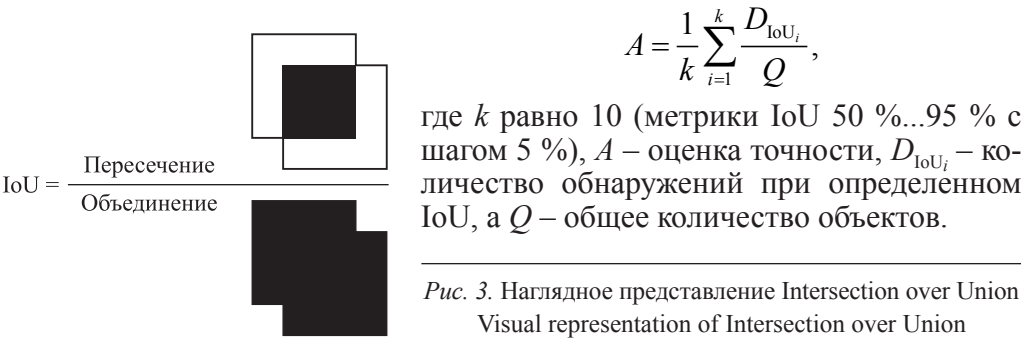


Рис. 3. Наглядное представление Intersection over Union
Visual representation of Intersection over Union

Другими словами, производится обнаружение объектов на оценочной выборке, затем определяется отношение количества обнаружений с IoU больше 50 % к количеству всех имеющихся объектов в оценочной выборке. Такое же отношение определяется для обнаружений с IoU больше 55 %, 60 и так далее до 95 %. Среднее арифметическое всех найденных отношений является усредненной оценкой точности модели обнаружения объектов.

Результаты экспериментов

Для определения наиболее подходящей модели обнаружения TensorFlow для решения задачи детектирования области глаз были проведены эксперименты с дообучением предварительно обученных моделей и обучением моделей с нуля.

Платформа TensorFlow содержит в себе функционал оценки точности работы обученной модели обнаружения. Далее представлены оценки точности для каждой модели, обученной с нуля и дообученной (см. таблицу).

Результаты экспериментов обучения Results of learning experiments

Название модели TensorFlow	Оценка точности	
	Обучение с нуля	Дообучение
CenterNet HourGlass104 512×512	0,671	0,456
CenterNet Resnet50 V1 FPN 512×512	0,525	0,337
CenterNet Resnet101 V1 FPN 512×512	0,544	0,389
CenterNet Resnet50 V2 512×512	0,513	0,297
CenterNet MobileNetV2 FPN 512×512	0,499	0,276
EfficientDet D0 512×512	0,527	0,371
SSD MobileNet v2 320×320	0,478	0,224
SSD MobileNet V1 FPN 640×640	0,511	0,292
SSD MobileNet V2 FPNLite 320×320	0,695	0,474
SSD ResNet50 V1 FPN 640×640 (RetinaNet50)	0,553	0,364
SSD ResNet101 V1 FPN 640×640 (RetinaNet101)	0,557	0,364
SSD ResNet152 V1 FPN 640×640 (RetinaNet152)	0,554	0,357
Faster R-CNN ResNet50 V1 640×640	0,521	0,336
Faster R-CNN ResNet101 V1 640×640	0,526	0,338
Faster R-CNN ResNet152 V1 640×640	0,534	0,375
Faster R-CNN Inception ResNet V2 640×640	0,583	0,377

В результате экспериментов модель «SSD MobileNet V2 FPNLite 320×320» получила самые высокие оценки точности после обучения с нуля и после дообучения. В случае обучения с нуля оценка точности составила 0,695, а в случае дообучения – 0,474.

Далее представлен график общих потерь (сумм потерь классификации и локализации, которые были записаны во время обучения с нуля) для этой модели (рис. 4).

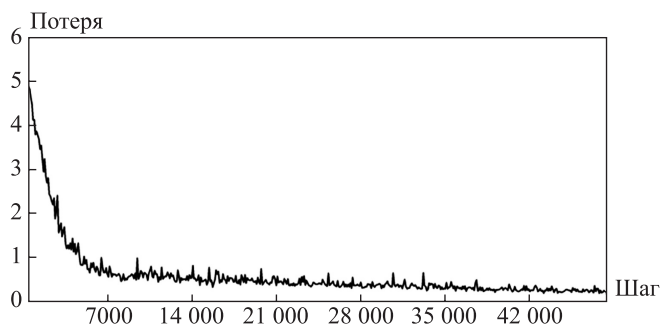


Рис. 4. График потерь
Graph of losses

Далее представлены значения отношений верно обнаруженных объектов в зависимости от различных IoU для обученной с нуля модели «SSD MobileNet V2 FPNLite 320×320» (рис. 5).

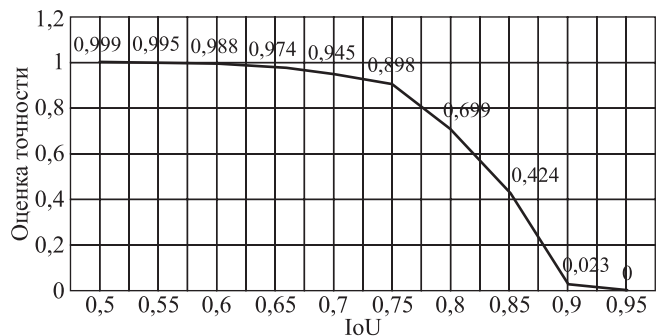


Рис. 5. Оценка точности в зависимости от IoU
Estimation of accuracy depending on IoU

Обученная с нуля модель «SSD MobileNet V2 FPNLite 320×320» может использоваться не только в информационной системе автоматического диагностирования косоглазия, разрабатываемой автором статьи, но и в других информационных системах, в которых требуется функционал обнаружения на фотографии областей с глазами. В зависимости от требований касательно точности обнаружения модель будет иметь различную точность в процентном соотношении. Например, если требования по точности в программе очень строгие – IoU должно быть больше 0,75, то максимальная точность обученной с нуля составит 89,8 %. А если требования по точности не строгие, например, минимальное значение IoU должно быть 0,5, то точность составит 99,9 % (рис. 5). Далее представлено наглядное сравнение различных значений IoU (рис. 6).

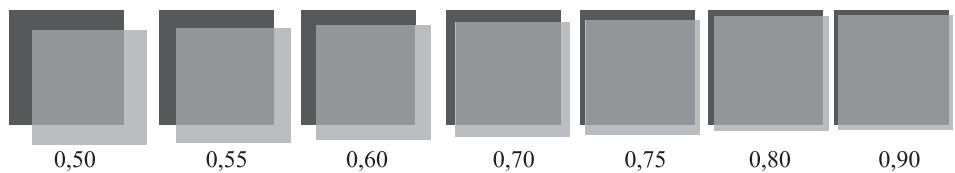


Рис. 6. Наглядное сравнение различных IoU
Visual comparison of various IoUs

Согласно исследованиям [6, 13, 16–18] значение IoU больше 0,5 обычно считается хорошим результатом обнаружения объекта.

Демонстрация работы модели

Для демонстрации работы обученной с нуля модели «SSD MobileNet V2 FPNLite 320×320» был написан скрипт на языке Python с использованием библиотеки TensorFlow. Результатом работы скрипта являются изображения с выделенными на нем областями с глазами (рис. 7).

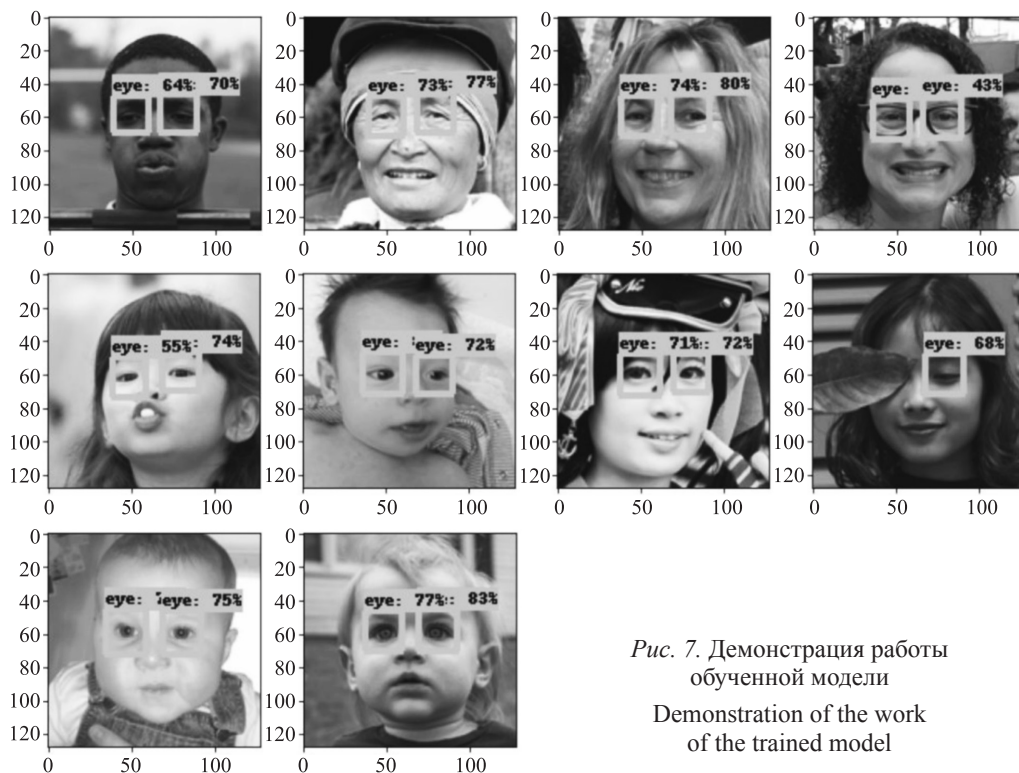


Рис. 7. Демонстрация работы обученной модели
Demonstration of the work of the trained model

Проценты на рисунке обозначают степень соответствия найденных объектов набору признаков модели, по которым осуществляется обнаружение. Данная демонстрация показывает, что модель успешно распознает глаза на фотографиях людей вне зависимости от окружения, головных уборов, причесок, возрастов, национальностей, эмоций, аксессуаров (украшений, очков) и помех.

Заключение

В статье представлены результаты анализа сравнительной эффективности моделей обнаружения TensorFlow при решении задачи обнаружения области глаз на фотографии лица человека. Проведенные эксперименты позволили установить, что для решения этой задачи больше всего подходит модель «SSD MobileNet V2 FPNLite 320×320», обученная с нуля, поскольку точность работы модели составляет 99,9 % для программ с требованием минимального IoU равным 0,5.

Результаты экспериментов могут быть использованы в различных исследованиях, в которых применяется платформа TensorFlow для обнаружения объектов на изображениях, и производится обнаружение объектов только одного класса.

Обученную модель предполагается использовать в информационной системе автоматического диагностирования косоглазия.

Список источников

1. *Бурдуковский С.О.* Анализ и оценка программ глубинного обучения // Развитие интеллектуального потенциала молодежи Кубани–2019: Материалы II Всероссийской научно-практической конф., 13–18 мая 2019 г. Анапа: Анапский филиал ФГБОУ ВО «МПГУ», 2019. С. 18–24.
2. *Совнелъ Д.С.* Анализ эмоций человека по портретному снимку // 56-я научная конференция аспирантов, магистрантов и студентов БГУИР, 18–20 мая 2020 г. Минск: БГУИР, 2020. С. 159–161.
3. *Araujo G.M., Ribeiro F.M.L., Silva E.A.B., Goldenstein S.K.* Fast eye localization without a face model using inner product detectors // 2014 IEEE International Conference on Image Processing, 27–30 october 2014. Paris, 2014. P. 1366–1370. DOI: 10.1109/ICIP.2014.7025273
4. *LeCun Y., Bengio Y., Hinton G.* Deep learning // Nature. 2015. Vol. 521. P. 436–444. ISSN 1476-4687. DOI: 10.1038/nature14539
5. *Markuš N., Frljak M., Pandžić I.S., Ahlberg J., Forchheimer R.* Eye pupil localization with an ensemble of randomized trees // Pattern Recognition. 2014. Vol. 47. P. 578–587. DOI: 10.1016/j.patcog.2013.08.008
6. *Maxwell A., Warner T.A., Guillén L.A.* Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies – Part 1: Literature Review // Remote Sensing. 2021. Vol. 13. P. 1–27. ISSN 2072-4292. DOI: 10.3390/rs13132450
7. *Rehman T.U., Mahmud Md.S., Chang Y.K., Shin J.J.J.* Current and future applications of statistical machine learning algorithms for agricultural machine vision systems // Computers and Electronics in Agriculture. 2019. Vol. 156. P. 585–605. ISSN 0168-1699. DOI: 10.1016/j.compag.2018.12.006
8. *Saha D., Manickavasagan A.* Machine learning techniques for analysis of hyper-spectral images to determine quality of food products: A review // Current Research in Food Science. 2020. Vol. 4. P. 28–44. ISSN 2665-9271. DOI: 10.1016/j.crfs.2021.01.002
9. *Swirski L., Bulling A., Dodgson N.A.* Robust real-time pupil tracking in highly off-axis images // Proceedings of the 7th Eye Tracking Research and Applications Symposium. 2012. P. 173–176. DOI: 10.1145/2168556.2168585
10. *Timm F., Barth E.* Accurate eye centre localisation by means of gradients // VISAPP 2011 – Proceedings of the Sixth International Conference on Computer Vision Theory and Applications, 5–7 march 2011. Vilamora, 2011. P. 125–130.
11. *Valenti R., Gevers T.* Accurate eye center location through invariant isocentric patterns // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2012. Vol. 34. P. 1785–1798. DOI: 10.1109/TPAMI.2011.251
12. *Xie D., Zhang L., Bai L.* Deep Learning in Visual Computing and Signal Processing // Applied Computational Intelligence and Soft Computing. 2017. Vol. 2017. P. 1–13. DOI: 10.1155/2017/1320780.
13. *Zhang X., Han L., Han L., Zhu L.* How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? // Remote Sensing. 2020. Vol. 12. P. 1–29. ISSN 2072-4292. DOI: 10.3390/rs12030417

14. COCO – Common Objects in Context. [Электронный ресурс]. URL: <https://cocodataset.org/> (дата обращения: 26.03.2022).
15. Flickr-Faces-HQ Dataset (FFHQ). [Электронный ресурс]. URL: <https://github.com/NVlabs/ffhq-dataset> (дата обращения: 08.03.2022).
16. Intersection over Union (IoU) for object detection – PyImageSearch. [Электронный ресурс]. URL: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> (дата обращения: 03.04.2022).
17. Measuring Labelling Quality with IOU and F1 Score / by Isaac Tan Supahands Tech Blog // Medium. [Электронный ресурс]. URL: <https://medium.com/supahands-techblog/measuring-labelling-quality-with-iou-and-f1-score-1717e29e492f> (дата обращения: 03.04.2022).
18. Object Detection -IOU-Intersection Over Union / by Ananth Sandeep // Medium. [Электронный ресурс]. URL: <https://medium.com/@nagsan16/object-detection-iou-intersection-over-union-73070cb11f6e> (дата обращения: 03.04.2022).
19. ResearchGate – Best percent of divided the training data and testing data. [Электронный ресурс]. URL: https://www.researchgate.net/post/can_someone_recommend_what_is_the_best_percent_of_divided_the_training_data_and_testing_data_in_neural_network_7525_or_8020_or_9010 (дата обращения: 08.03.2022).
20. ResearchGate – Is there an ideal ratio between a training set and validation set. [Электронный ресурс]. URL: <https://www.researchgate.net/post/Is-there-an-ideal-ratio-between-a-training-set-and-validation-set-Which-trade-off-would-you-suggest> (дата обращения: 08.03.2022).
21. TensorFlow. [Электронный ресурс]. URL: <https://www.tensorflow.org> (дата обращения: 20.03.2022).

References

1. Burdukovskij S.O. Analiz i ocenka programm glubinnogo obuchenija [Analysis and evaluation of deep learning programs]. Razvitie intellektual'nogo potencijala molodezhi Kubani – 2019: Materialy II Vserossijskoj nauchno-prakticheskoy konf., 13–18 maja 2019 g. Anapa: Anapskij filial FGBOU VO «MPGU», 2019. P. 18–24.
2. Sovpel' D.S. Analiz jemocij cheloveka po portretnomu snimku [Analysis of human emotions in a portrait photograph]. 56-ja nauchnaja konferencija aspirantov, magistrantov i studentov BGUIR, 18–20 maja 2020 g. Minsk: BGUIR, 2020. P. 159–161.
3. Araujo G.M, Ribeiro F.M.L., Silva E.A.B., Goldenstein S.K. Fast eye localization without a face model using inner product detectors // 2014 IEEE International Conference on Image Processing, 27–30 october 2014. Paris, 2014. P. 1366–1370. DOI: 10.1109/ICIP.2014.7025273
4. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. Vol. 521. P. 436–444. ISSN 1476-4687. DOI: 10.1038/nature14539
5. Markuš N., Frljak M., Pandžić I.S., Ahlberg J., Forchheimer R. Eye pupil localization with an ensemble of randomized trees // Pattern Recognition. 2014. Vol. 47. P. 578–587. DOI: 10.1016/j.patcog.2013.08.008
6. Maxwell A., Warner T.A., Guillén L.A. Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies – Part 1: Literature Review // Remote Sensing. 2021. Vol. 13. P. 1–27. ISSN 2072-4292. DOI: 10.3390/rs13132450
7. Rehman T.U., Mahmud Md.S., Chang Y.K., Shin J.J.J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems // Computers and Electronics in Agriculture. 2019. Vol. 156. P. 585–605. ISSN 0168-1699. DOI: 10.1016/j.compag.2018.12.006
8. Saha D., Manickavasagan A. Machine learning techniques for analysis of hyper-spectral images to determine quality of food products: A review // Current Research in Food Science. 2020. Vol. 4. P. 28–44. ISSN 2665-9271. DOI: 10.1016/j.crfs.2021.01.002

9. Swirski L., Bulling A., Dodgson N.A. Robust real-time pupil tracking in highly of-axis images // *Proceedings of the 7th Eye Tracking Research and Applications Symposium*. 2012. P. 173–176. DOI: 10.1145/2168556.2168585
10. Timm F., Barth E. Accurate eye centre localisation by means of gradients // *VISAPP 2011 – Proceedings of the Sixth International Conference on Computer Vision Theory and Applications*, 5–7 march 2011. Vilamora, 2011. P. 125–130.
11. Valenti R., Gevers T. Accurate eye center location through invariant isocentric patterns // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012. Vol. 34. P. 1785–1798. DOI: 10.1109/TPAMI.2011.251
12. Xie D., Zhang L., Bai L. Deep Learning in Visual Computing and Signal Processing // *Applied Computational Intelligence and Soft Computing*. 2017. Vol. 2017. P. 1–13. DOI: 10.1155/2017/1320780.
13. Zhang X., Han L., Han L., Zhu L. How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? // *Remote Sensing*. 2020. Vol. 12. P. 1–29. ISSN 2072-4292. DOI: 10.3390/rs12030417
14. COCO – Common Objects in Context. [Electronic resource]. Available at: <https://cocodataset.org/> (accessed: 26.03.2022).
15. Flickr-Faces-HQ Dataset (FFHQ). [Electronic resource]. Available at: <https://github.com/NVlabs/ffhq-dataset> (accessed: 08.03.2022).
16. Intersection over Union (IoU) for object detection – PyImageSearch. [Electronic resource]. Available at: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> (accessed: 03.04.2022).
17. Measuring Labelling Quality with IOU and F1 Score / by Isaac Tan Supahands Tech Blog // Medium. [Electronic resource]. Available at: <https://medium.com/supahands-techblog/measuring-labelling-quality-with-iou-and-f1-score-1717e29e492f> (accessed: 03.04.2022).
18. Object Detection -IOU-Intersection Over Union / by Ananth Sandeep // Medium. [Electronic resource]. Available at: <https://medium.com/@nagsan16/object-detection-iou-intersection-over-union-73070cb11f6e> (accessed: 03.04.2022).
19. ResearchGate – Best percent of divided the training data and testing data. [Electronic resource]. Available at: https://www.researchgate.net/post/can_someone_recommend_what_is_the_best_percent_of_divided_the_training_data_and_testing_data_in_neural_network_7525_or_8020_or_9010 (accessed: 08.03.2022).
20. ResearchGate – Is there an ideal ratio between a training set and validation set. [Electronic resource]. Available at: <https://www.researchgate.net/post/Is-there-an-ideal-ratio-between-a-training-set-and-validation-set-Which-trade-off-would-you-suggest> (accessed: 08.03.2022).
21. TensorFlow. [Electronic resource]. Available at: <https://www.tensorflow.org> (accessed: 20.03.2022).

Сведения об авторе:

С.О. Бурдуковский – аспирант, Новосибирский государственный университет экономики и управления «НИНХ», Новосибирск, Российская Федерация.

Information about the author:

S.O. Burdukowsky – Post-Graduate Student, Novosibirsk State University of Economics and Management, Novosibirsk, Russian Federation.

<i>Статья поступила в редакцию</i>	<i>10.03.2022</i>	<i>The article was submitted</i>	<i>10.03.2022</i>
<i>Одобрена после рецензирования</i>	<i>11.04.2022</i>	<i>Approved after reviewing</i>	<i>11.04.2022</i>
<i>Принята к публикации</i>	<i>11.05.2022</i>	<i>Accepted for publication</i>	<i>11.05.2022</i>