

РАЗРАБОТКА ТЕХНОЛОГИИ ОПРЕДЕЛЕНИЯ ПОРОДЫ ДЕРЕВЬЕВ С ПРИМЕНЕНИЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Д.Ю. Войтов¹✉, С.Б. Васильев², Д.В. Кормилицын²

¹ООО «Инжиниринговый центр «Автоматика и робототехника МГТУ им. Н.Э. Баумана», 105005, г. Москва, ул. 2-я Бауманская, д. 5, стр. 1

²Мытищинский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)», 141005, Россия, Московская обл., г. Мытищи, ул. 1-я Институтская, д. 1

leonickss@yandex.ru

Представлена разработанная технология определения на фотографии породы деревьев березы повислой (*Betula pendula* Roth), а также материалы изучения различных известных нейронных сетей классификаторов с определением объектов. Среди них выбрана нейронная сеть-классификатор YOLOv4, как самая перспективная для дальнейшего развития технологии. Изучен механизм разметки фотоизображений для формирования примеров обучения. Сформированы методика разметки на изображениях, два различных датасета для дообучения сети. Проведено алгоритмическое увеличение датасета путем трансформации фотоизображений и наложения фильтров. Определено различие в работе классификатора. Точность определения породы при обучении исключительно на фотоизображениях, содержащих березу повислую, составила 35 %, при обучении на датасете, содержащем другие деревья, — 71 %, на всем датасете — 75 %. Для демонстрации работы были определены деревья березы на фотографиях, выполненных в дендропарке Мытищинского филиала МГТУ им. Н.Э. Баумана. Для усовершенствования технологии рекомендуется дообучение сети для определения остальных пород деревьев. Технологию можно использовать для осуществления таксации конкретных пород деревьев, формирования размеченных датасетов для дальнейших разработок, как первичный элемент в системе анализа изображений деревьев, для исключения сторонних объектов на исходном фотоизображении.

Ключевые слова: компьютерное зрение, береза повислая, таксация, YOLO, нейронные сети

Ссылка для цитирования: Войтов Д.Ю., Васильев С.Б., Кормилицын Д.В. Разработка технологии определения породы деревьев с применением компьютерного зрения // Лесной вестник / Forestry Bulletin, 2023. Т. 27. № 1. С. 60–66. DOI: 10.18698/2542-1468-2023-1-60-66

В современном высокотехнологичном мире все чаще возникает вопрос о замене различными алгоритмами человеческого труда. Лесная отрасль не является исключением. Нахождение и обнаружение объектов представляет собой бессознательное свойство человеческого глаза, что позволяет решать широкий спектр задач. Развитие современных технологий дает возможность подменить человека машиной. Область знаний, отвечающая за распознавание объектов на фотоизображениях, называется компьютерным зрением.

Цель работы

Цель работы — рассмотрение технологии определения объектов на фотоизображении.

Материалы и методы

Определяемым объектом принят самый распространенный вид дерева — береза повислая (*Betula pendula* Roth). Основная задача исследования — формирование метода определения участков фотоизображения, содержащих *Betula pendula* Roth, выявление механизмов оценки эффективности разработанной технологии, изучение перспектив-

ности и программная реализация. Данную задачу можно отнести к типовым, решаемым с помощью технологии компьютерного зрения.

Обработка фотоизображений для компьютера нетривиальная задача [1]. У объектов на фотоизображении часто не видно четких границ, одни объекты могут перекрывать другие и довольно сильно отличаются один от другого. Таким образом, использовать стандартные алгоритмы и решения для определения границ объектов и их типов не всегда приемлемо. Для решения возникшей проблемы разработаны и широко применяются технологии компьютерного зрения.

Одним из самых распространенных инструментов в компьютерном зрении являются нейросетевые технологии [2]. Их суть заключается в формировании определенной сетевой структуры, к частным элементам которой относится упрощенная модель нейронов [3], при прохождении информации по которой формируется решение задачи.

Нейронные сети определения конкретных объектов называются определителем объектов (object detector) [4] и представляют собой довольно сложные для понимания структуры, состоящие из десятков миллионов нейронов. Нейронные сети предварительно обучаются на больших выборках

различных данных, включающих в себя сотни тысяч размеченных фотоизображений. Тем не менее, для специализации нейронных сетей необходимы существенно меньшие объемы данных, что позволяет подстраивать этот инструмент под себя.

По методике обработки фотоизображений нейронные сети можно условно подразделить на два типа: одно- и двухэтапные [5].

Двухэтапные сети предварительно разбивают изображения на регионы, в которых отмечаются области, с наибольшей вероятностью содержащие объекты. После чего такие области подвергаются дополнительному рассмотрению с помощью основного классификатора, который более точно определяет границы объектов и их относимость к какому-либо классу.

Одноэтапные сети генерируют некоторое количество рамок на исходном фотоизображении, после чего доопределяют их расположение и границы. Этот метод позволяет существенно ускорить процесс определения объекта [5].

Для определения точности обычно используется метрика отношения площадей окружающих рамок [6]. Ее суть заключается в соотношении пересекающихся областей предполагаемых алгоритмом границ вокруг объекта и предварительно размеченной истинной рамкой, соответствующей ему. Таким образом, результат считается тем лучше, чем он ближе к единице.

Для задач с большим количеством различных категорий определяемых объектов применяется также метрика усредненной по всем классам точности [7]. Это позволяет получать менее узконаправленные классификаторы. В настоящей работе такая метрика не рассматривается, однако она может быть важной при дальнейшем развитии технологии.

Третьей метрикой [4] оценки систем называют сложность обработки фотоизображения и разметки на нем всех объектов. Несмотря на то, что ее влияние считается незначительным в решении задач определения объектов на фотографиях, низкая скорость выполнения алгоритмов и высокая ресурсоемкость удлиняют время обработки изображений.

Таким образом, нельзя склоняться к исключительно точным алгоритмам, поскольку это может привести к значительным ограничениям в масштабировании и дальнейшем развитии технологии.

Результаты и обсуждение

Среди огромного разнообразия различных определителей можно выделить несколько по-настоящему используемых и популярных, которые опережают остальные по всем параметрам и конкурируют лишь между собой. К ним относятся: RCNN, RFCN, SSD, YOLOv4 [8].

Первые две относятся к двухэтапным сетям, последующие — к одноэтапным.

RCNN (Region — Based Convolutional Neural Network) [8] — прямой представитель двухэтапной сети. Суть алгоритма определения фотоизображений состоит в определении зон интереса и последующей классификации объектов в них. Алгоритм показывает точность классификации около 70,4 %, а время обработки фотоизображений составило в среднем более 200 мс [9, 10].

RFCN (Region — Based Fully-Convolutional Network) [11] — собой ускоренный вариант RCNN. Одновременно выполняет работу по классификации и построению зон интереса. Полученные результаты совмещаются для корректировки итоговых границ объектов. Имеет точность до 80,5 %, при времени обработки около 85 мс [9, 10].

SSD (Single — Shot Detector) и YOLOv4 (You Only Look Once) [12, 13] отличаются исключительно по структуре нейронной сети, сохраняя единый подход. SSD показывает точность около 75,8 % с временем обработки 61 мс, а YOLOv4 — 79,6 % и 29 мс соответственно [9, 10].

На основании представленных характеристик было принято решение о использовании нейронной сети YOLOv4, поскольку при незначительно меньшем по сравнению с RFCN качестве определения (около 1 %), время обработки фотоизображения уменьшилось практически в 3 раза, т. е. YOLOv4 можно использовать для систем реального времени (рис. 1) [14].

Для дообучения нейронной сети YOLOv4 было собрано несколько сотен различных фотоизображений, содержащих березу повислую (*Betula pendula* Roth). Поскольку задача заключалась в определении березы на фотоизображениях, возник вопрос о его разметке: почти на половине фотоизображений дерево не было представлено в полную величину, вследствие чего YOLOv4 могла неправильно распознать дерево или не найти его на фотоизображении. Для того, чтобы решить проблему кроны и ствол дерева разделили на два типа: кроны выделили в класс «кроны» (1-й), а стволы — в класс «березы» (2-й). Затем был написан скрипт на языке python, который перед формированием итогового изображения с разметкой находил пересекающиеся ограничительные рамки обоих классов и объединял их [15]. Остальные объекты класса «крона» удалялись. Кроме этого, данное решение позволяло получить несколько составных элементов дерева, что стало полезно для изучения. После принятия решения о способе разметки в ручном режиме с использованием специального программного обеспечения (ПО) Label Image [16] были определены кроны и стволы. Важно отметить, что отмечались кроны, принадлежащие и к другим породам деревьев.

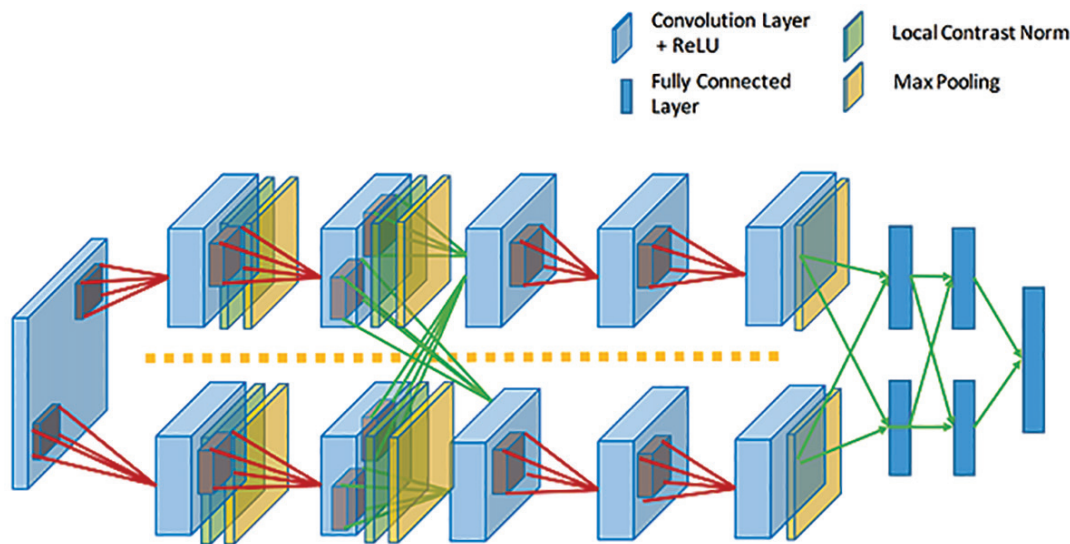


Рис. 1. Структура нейронной сети YOLOv4 [14]
 Fig. 1. The structure of the YOLOv4 neural network [14]

Сама разметка представляла собой выделение мышью прямоугольника на изображении, содержащем необходимые объекты, после чего им присваивался соответствующий класс (1-й или 2-й). После этого программа Label.Image самостоятельно создает дополнительную папку, в которой хранятся текстовые файлы [16], соответствующие изображениям. Структура файла [17, 18] представляет собой пять чисел, написанных через запятую, для каждого из отмеченных на фотографии объектов: первое число — номер класса объекта; второе и третье — координаты X и Y левой нижней точки выделенного на изображении прямоугольника в виде нормализованного значения (отношение номера пикселя к общему количеству пикселей); четвертое и пятое — координаты правой верхней точки. Это позволяет осуществить разметку в автоматическом режиме при программном известном изменении исходных фотоизображений.

Количество изображений искусственно увеличили в 2 раза путем добавления различных шумовых фильтров и небольших случайных поворотов исходных изображений [19, 20]. После этого были созданы файлы с разметкой, в которых проведено программное вычисление новых границ разметки. В отношении некоторых из случайно сгенерированных фотографий выполнили дополнительную проверку во избежание ошибок и в целях достижения точности полученного алгоритма. Для изучения влияния датасета подобранные изображения условно разделили на две группы: изображения, содержащие исключительно березу повислую (*Betula pendula* Roth) в единственном экземпляре, и изображения, на которых находились и другие породы деревьев.

В каждой из групп выделили тренировочное и тестовое подмножества, поскольку выделение валидационного подмножества было определено как нецелесообразное в связи с недостаточным количеством изображений и отсутствия необходимости доработки архитектуры сети [21].

Таким образом, на тренировочную выборку было выделено 70 % всех изображений, а на тестовую — 30 %. Распределение проводилось программно, случайным образом. После этого программно было проведено масштабирование всех фотоизображений до единого размера 640×640 пикселей [22].

Результаты точности определения березы повислой (*Betula pendula* Roth) на тестовом датасете в зависимости от обучающей выборки

The results of the accuracy of determining *Betula pendula* Roth on the test dataset, depending on training sample

Содержание датасета	Точность определения, %
Исключительно береза повислая (<i>Betula pendula</i> Roth)	35
Множественные породы деревьев	71
Совмещенный	75

В каждую эпоху обучения велось сохранение весов в сети для возможности выбора оптимальной с точки зрения точности на тестовом подмножестве (таблица).

При обучении на фотографиях, содержащих исключительно березу повислую, точность на тестовом подмножестве оказалась значительно ниже, поскольку сеть начинает маркировать как березу каждый объект, похожий на дерево. При

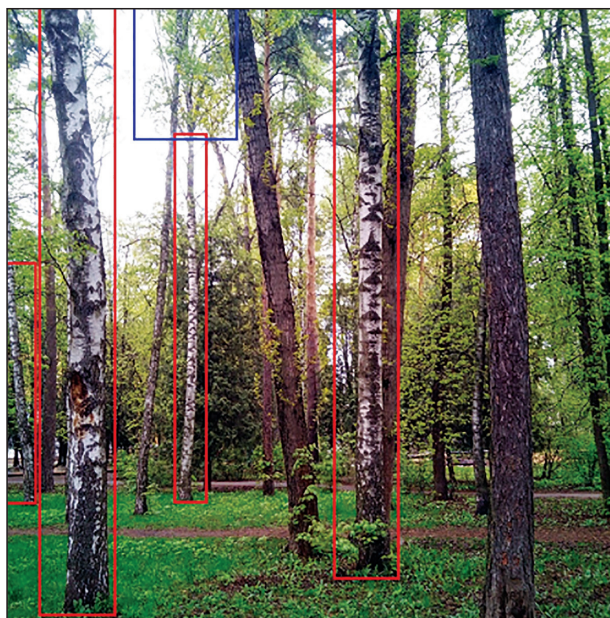


Рис. 2. Пример определения на изображении дерева породы береза повислая

Fig. 2. An example of determining on the image of a European white birch tree species

этом при обучении на фотографиях, содержащих не только березу, нейронная сеть не всегда точно определяла это дерево.

Следовательно, важны фотоизображения обоих характеров, причем изображения с присутствием иных пород еще важнее. Отметим, что точность определения повышается с увеличением размера датасета.

В целях демонстрации результатов дообученная на всем датасете модель была использована для определения деревьев березы на фотографиях, выполненных в дендропарке МГТУ им. Н.Э. Баумана (Мытищинский филиал) (рис. 2).

Выводы

Рассмотренная технология может использоваться в широком спектре различных задач [23, 24], в частности при таксации лесных массивов [25] с дообучением модели на предмет других пород деревьев. Кроме того, технологию можно применить для выделения областей с деревом березы для дальнейших обработок изображений [26, 27]. Технология может служить для пополнения общих датасетов в сообществе разработчиков, для первичной обработки в системах, целью которых является выявление фотоизображения с деревом березы для получения желаемых результатов, для фильтрации среды посторонних объектов.

К тому же возможности технологии позволяют осуществлять обработку в режиме реального времени [28–30] и решать задачи с крайне незначительными доработками.

Список литературы

- [1] RoboCraft. SLAM. URL: <http://robocraft.ru/blog/technology/724.html> (дата обращения 30.04.2022).
- [2] Review: YOLOv1 – You Only Look Once (Object Detection). URL: <https://towardsdatascience.com/yolov1-you-only-look-once-object-detection-e1f3ffec8a89> (дата обращения 21.04.2022).
- [3] Batch-normalization. URL: <http://neerc.ifmo.ru/wiki/index.php?title=Batch-normalization> (дата обращения 10.05.2022).
- [4] Redmon J., Farhadi A. YOLOv3: An Incremental Improvement, 2018. URL: <https://arxiv.org/abs/1804.02767> (дата обращения 30.04.2022).
- [5] Convex Hull, (Minimum) Bounding Box, and Minimum Enclosing Circle. URL: <http://dwwoll.de/rexrepos/posts/diagBounding.html> (дата обращения 10.05.2022).
- [6] YOLO: Real-Time Object Detection. URL: <https://pjreddie.com/darknet/yolo/> (дата обращения 30.04.2022).
- [7] Szegedy C., Lofe S., Vanhoucke V., Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 2016. URL: <https://arxiv.org/abs/1602.07261v2>
- [8] Chowdhury A.P., Kulkarni P., Bojnordi M.N. MB-CNN: Memristive binary convolutional neural networks for embedded mobile devices // J. of Low Power Electronics and Applications, 2018, no. 8(4), p. 38. DOI: 10.3390/JLPEA8040038
- [9] Neural Processing Unit. URL: <https://www.pcmag.com/encyclopedia/term/neural-processing-unit> (дата обращения 22.04.2022).
- [10] Arakawa Y., Matsuda Y. Gamification Mechanism for Enhancing a Participatory Urban Sensing: Survey and Practical Results // J. of Information Processing, 2016, no. 24(1), pp. 31–38. DOI:10.2197/ipsjip.24.31
- [11] Yasumoto K., Yamaguchi H., Shigeno H. Survey of real-time processing technologies of iot data streams // J. of Information Processing, 2016, no. 24(2), pp. 95–202.
- [12] Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G.S., Davis A., Dean J., Devin M. Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016. URL: <https://doi.org/10.48550/arXiv.1603.04467> (дата обращения 11.05.2022).
- [13] Piccardi M. Background subtraction techniques: a review // Conference: Systems, Man and Cybernetics IEEE International Conference on Volume: 4, 2004. DOI:10.1109/ICSMC.2004.1400815
- [14] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // February 2001. Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, no. 1, pp. I–511 – I–518, vol. 1. DOI:10.1109/CVPR.2001.990517
- [15] Team O. Opencv. URL: <https://opencv.org> (дата обращения 22.04.2022).
- [16] LeCun Y., Haffner P., Bottou L., Bengio Y. Object recognition with gradient-based learning. Palermo, Sicily: Springer Verlag, 1999, pp. 319–345
- [17] Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks Advances//Neural Information Processing, 2012, no. 25(2). DOI:10.1145/3065386
- [18] Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciampi F., Ghafoorian M., Van Der Laak J.A., Van Ginneken B., Sanchez C.I. A survey on deep learning in medical image analysis // Medical Image Analysis, 2017, no. 42, pp. 60–88.

- [19] Han J., Zhang D., Cheng G., Liu N., Xu D. Advanced deep learning techniques for salient and category-specific object detection: a survey // *IEEE Signal Processing Magazine*, 2018, pp. 84–100. DOI:10.1109/MSP.2017.2749125
- [20] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition, 2014. <https://doi.org/10.48550/arXiv.1409.1556>
- [21] Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions // *Conference: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. DOI:10.1109/CVPR.2015.7298594
- [22] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A.C. Ssd: Single shot multibox detector // *Conference: European Conference on Computer Vision*, 2016. DOI:10.1007/978-3-319-46448-0_2
- [23] Howard A.G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017. <https://doi.org/10.48550/arXiv.1704.04861>
- [24] Redmon J., Farhadi A. Yolov3: An incremental improvement, 2018. <https://doi.org/10.48550/arXiv.1804.02767>
- [25] Redmon J., Divvala S., Girshick R., Farhadi A. You only look once: Unified, real-time object detection // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016. <https://doi.org/10.48550/arXiv.1506.02640>
- [26] Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. Simple online and realtime tracking // *Image Processing (ICIP), IEEE International Conference*, 2016, pp. 3464–3468. DOI:10.1109/ICIP.2016.7533003
- [27] Kalman R.E. A new approach to linear filtering and prediction problems // *J. of Basic Engineering (American Society of Mechanical Engineers)*, 1960, vol. 82, iss. 1, pp. 35–45. DOI:10.1115/1.3662552
- [28] Kuhn H.W. The hungarian method for the assignment problem // *Naval research logistics quarterly*, 1955, no. 2(1), pp. 83–97. <http://dx.doi.org/10.1002/nav.3800020109>
- [29] Lazorenko A. Tensorflow performance test: Cpu vs gpu, 2017. URL: <https://medium.com/@andriylazorenko/tensorflow-performance-test-cpu-vs-gpu-79fcd39170c> (дата обращения 10.04.2022).
- [30] Ouaknine A. Review of Deep Learning Algorithms for Object Detection, 2018. URL: <https://medium.com/zylapp/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852> (дата обращения 21.04.2022).

Сведения об авторах

Войтов Даниил Юрьевич✉ — инженер ИЦ «АР» МГТУ им. Н.Э. Баумана, leonickss@yandex.ru
Васильев Сергей Борисович — канд. с.-х. наук, доцент, зав. кафедрой лесных культур, селекции и дендрологии ФГБОУ ВО «МГТУ им. Н.Э. Баумана (Мытищинский филиал)», svasilyev@mgul.ac.ru
Кормилицын Дмитрий Владимирович — аспирант кафедры лесных культур, селекции и дендрологии ФГБОУ ВО «МГТУ им. Н.Э. Баумана (Мытищинский филиал)», gurkordmi@gmail.com

Поступила в редакцию 28.06.2022.

Одобрено после рецензирования 13.09.2022.

Принята к публикации 16.11.2022.

TECHNOLOGY DEVELOPMENT FOR DETERMINING TREE SPECIES USING COMPUTER VISION

D.Y. Voytov¹✉, S.B. Vasil'ev², D.V. Kormilitsyn²

¹Engineering Center «Automation and Robotics» BMSTU, 5, build. 1, 2nd Baumanskaya st., 105005, Moscow, Russia

²BMSTU (Mytishchi branch), 1, 1st Institutskaya st., 141005, Mytishchi, Moscow reg., Russia

leonickss@yandex.ru

A technology has been developed to determine the European white birch (*Betula pendula* Roth.) species in the photo. The differences of the known neural networks of classifiers with the definition of objects are studied. YOLOv4 was chosen as the most promising for further development of the technology. The mechanism of image markup for the formation of training examples has been studied. The method of marking on the image has been formed. Two different datasets have been formed to retrain the network. An algorithmic increase in the dataset was carried out by transforming images and applying filters. The difference in the results of the classifier is determined. The accuracy when training exclusively on images containing hanging birch was 35 %, the accuracy when training on a dataset containing other trees was 71 %, the accuracy when training on the entire dataset was 75 %. To demonstrate the work, birch trees were identified in photographs taken in the arboretum of the MF Bauman Moscow State Technical University. To improve the technology, additional training is recommended to determine the remaining tree species. The technology can be used for the implementation of taxation of specific tree species; the formation of marked datasets for further development; the primary element in the tree image analysis system, to exclude third-party objects in the original image.

Keywords: computer vision, hanging birch, taxation, YOLO, neural networks

Suggested citation: Voytov D.Yu., Vasil'ev S.B., Kormilitsyn D.V. *Razrabotka tekhnologii opredeleniya porody derev'ev s primeneniem komp'yuternogo zreniya development of technology for determining tree species using computer vision* [Technology development for determining tree species using computer vision]. *Lesnoy vestnik / Forestry Bulletin*, 2023, vol. 27, no. 1, pp. 60–66. DOI: 10.18698/2542-1468-2023-1-60-66

References

- [1] RoboCraft. SLAM. Available at: <http://robocraft.ru/blog/technology/724.html> (accessed 30.04.2022).
- [2] Review: YOLOv1 – You Only Look Once (Object Detection). Available at: <https://towardsdatascience.com/yolov1-you-only-look-once-object-detection-e1f3ffec8a89> (accessed 21.04.2022).
- [3] Batch-normalization. Available at: <http://neerc.ifmo.ru/wiki/index.php?title=Batch-normalization> (accessed 10.05.2022).
- [4] Redmon J., Farhadi A. YOLOv3: An Incremental Improvement, 2018. Available at: <https://arxiv.org/abs/1804.02767> (accessed 30.04.2022).
- [5] Convex Hull, (Minimum) Bounding Box, and Minimum Enclosing Circle. Available at: <http://dwoell.de/rexrepos/posts/diagBounding.html> (accessed 10.05.2022).
- [6] YOLO: Real-Time Object Detection. Available at: <https://pjreddie.com/darknet/yolo/> (accessed 30.04.2022).
- [7] Szegedy C., Loffe S., Vanhoucke V., Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 2016. Available at: <https://arxiv.org/abs/1602.07261v2>
- [8] Chowdhury A.P., Kulkarni P., Bojnordi M.N. MB-CNN: Memristive binary convolutional neural networks for embedded mobile devices. *J. of Low Power Electronics and Applications*, 2018, no. 8(4), p. 38. DOI: 10.3390/JLPEA8040038
- [9] Neural Processing Unit. Available at: <https://www.pcmag.com/encyclopedia/term/neural-processing-unit> (accessed 22.04.2022).
- [10] Arakawa Y., Matsuda Y. Gamification Mechanism for Enhancing a Participatory Urban Sensing: Survey and Practical Results. *J. of Information Processing*, 2016, no. 24(1), pp. 31–38. DOI:10.2197/ipsjip.24.31
- [11] Yasumoto K., Yamaguchi H., Shigeno H. Survey of real-time processing technologies of iot data streams. *J. of Information Processing*, 2016, no. 24(2), pp. 95–202.
- [12] Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G.S., Davis A., Dean J., Devin M. Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016. <https://doi.org/10.48550/arXiv.1603.04467> (accessed 11.05.2022).
- [13] Piccardi M. Background subtraction techniques: a review. Conference: Systems, Man and Cybernetics IEEE International Conference on Volume: 4, 2004. DOI:10.1109/ICSMC.2004.1400815
- [14] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. February 2001. Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, no. 1, pp. 1–511 – 1–518 vol. 1. DOI:10.1109/CVPR.2001.990517
- [15] Team O. Opencv. Available at: <https://opencv.org> (accessed 22.04.2022).
- [16] LeCun Y., Haffner P., Bottou L., Bengio Y. Object recognition with gradient-based learning. Palermo, Sicily: Springer Verlag, 1999, pp. 319–345
- [17] Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks *Advances // Neural Information Processing*, 2012, no. 25(2). DOI:10.1145/3065386
- [18] Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciampi F., Ghahfarooq M., Van Der Laak J.A., Van Ginneken B., Sanchez C.I. A survey on deep learning in medical image analysis // *Medical Image Analysis*, 2017, no. 42, pp. 60–88.
- [19] Han J., Zhang D., Cheng G., Liu N., Xu D. Advanced deeplearning techniques for salient and category-specific object detection: a survey // *IEEE Signal Processing Magazine*, 2018, pp. 84–100. DOI:10.1109/MSP.2017.2749125

- [20] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition, 2014. <https://doi.org/10.48550/arXiv.1409.1556>
- [21] Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions. Conference: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. DOI:10.1109/CVPR.2015.7298594
- [22] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A.C. Ssd: Single shot multibox detector. Conference: European Conference on Computer Vision, 2016. DOI:10.1007/978-3-319-46448-0_2
- [23] Howard A.G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017. <https://doi.org/10.48550/arXiv.1704.04861>
- [24] Redmon J., Farhadi A. Yolov3: An incremental improvement, 2018. <https://doi.org/10.48550/arXiv.1804.02767>
- [25] Redmon J., Divvala S., Girshick R., Farhadi A. You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. <https://doi.org/10.48550/arXiv.1506.02640>
- [26] Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. Simple online and realtime tracking. Image Processing (ICIP), IEEE International Conference, 2016, pp. 3464–3468. DOI:10.1109/ICIP.2016.7533003
- [27] Kalman R.E. A new approach to linear filtering and prediction problems. J. of Basic Engineering (American Society of Mechanical Engineers), 1960, vol. 82, iss. 1, pp. 35–45. DOI:10.1115/1.3662552
- [28] Kuhn H.W. The hungarian method for the assignment problem. Naval research logistics quarterly, 1955, no. 2(1), pp. 83–97. <http://dx.doi.org/10.1002/nav.3800020109>
- [29] Lazorenko A. Tensorflow performance test: Cpu vs gpu, 2017. Available at: <https://medium.com/@andriylazorenko/tensorflow-performance-test-cpu-vs-gpu-79fcd39170c> (accessed 10.04.2022).
- [30] Ouaknine A. Review of Deep Learning Algorithms for Object Detection, 2018. Available at: <https://medium.com/zylapp/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852> (accessed 21.04.2022).

Authors' information

Voytov Daniil Yur'evich ✉ — Engineer, Engineering Center «Automation and Robotics» BMSTU, leonickss@yandex.ru

Vasil'ev Sergey Borisovich — Cand. Sci. (Agriculture), Associate Professor of the BMSTU (Mytishchi branch), svasilyev@mgul.ac.ru

Kormilitsyn Dmitriy Vladimirovich — pg. of the BMSTU (Mytishchi branch), gurkordmi@gmail.com

Received 28.06.2022.

Approved after review 13.09.2022.

Accepted for publication 16.11.2022.

Вклад авторов: все авторы в равной доле участвовали в написании статьи
Авторы заявляют об отсутствии конфликта интересов
Authors' Contribution: All authors contributed equally to the writing of the article
The authors declare that there is no conflict of interest