

УДК 004.93

DOI: 10.22213/2410-9304-2022-2-61-67

Применение нейронной сети YOLOv5 для распознавания наличия средств индивидуальной защиты

С. А. Филичкин, магистрант, ИжГТУ имени М.Т. Калашникова, Ижевск, Россия

С. В. Вологдин, доктор технических наук, доцент, ИжГТУ имени М.Т. Калашникова, Ижевск, Россия

В статье рассматриваются вопросы применения нейронной сети YOLO для автоматического обнаружения и идентификации средств индивидуальной защиты у персонала предприятий на цифровых изображениях. Приведен сравнительный анализ различных архитектур сети YOLOv5 в зависимости от скорости обучения нейронной сети и точности распознавания. Рассмотрены инструменты для обучения нейронной сети на языке Python в среде Google Colab, которая позволяет эффективно использовать удаленные вычислительные мощности графических процессоров.

Рассмотрен пример обучения сети на основе собранного датасета с шестью классами для распознавания наличия средств индивидуальной защиты (с маской; без маски; маска надета неправильно; с головным убором; без головного убора; спец. одежда). Описаны настройки нейронной сети YOLOv5 и рекомендации по улучшению распознавания объектов посредством дополнения датасета изображениями с определенными условиями. Произведены 4 варианта обучения сети с применением архитектуры YOLOv5s и YOLOv5l, а также с различным количеством эпох. Подобрана оптимальная архитектура нейронной сети и количество необходимых эпох. Рассмотрено понятие средней точности (Mean Average Precision – mAP) и приведен сравнительный анализ точности предсказания от выбранной архитектуры сети и настроек.

Приведен практический пример распознавания наличия средств индивидуальной защиты на цифровых изображениях с помощью обученной нейронной сети YOLOv5. В среднем точность распознавания составляет 0,7, что является неплохим результатом для решения данного класса задач. В работе рассматриваются вопросы по повышению точности распознавания объектов. Приводятся планы по обнаружению и детекции новых классов средств индивидуальной защиты, а также подтверждение факта прохождения регламентных процедур персоналом предприятий на основе цифрового видеоряда.

Ключевые слова: компьютерное зрение, искусственный интеллект, распознавание образов, нейронная сеть, разметка изображений, YOLO.

Введение

Обеспечение персонала предприятий средствами индивидуальной защиты важно не только для охраны здоровья людей, но и для соблюдения производственных регламентов выполнения работ. Особенно данная тематика актуальна сейчас в период пандемии COVID-2019.

Целью данного исследования является повышение точности распознавания наличия средств индивидуальной защиты у персонала предприятий. Данная задача распознавания образов является задачей классификации в машинном обучении и относится к разделу обучения с учителем. В статье рассматриваются вопросы распознавания средств индивидуальной защиты с помощью нейронной сети YOLO (You Only Look Once).

Как известно, нейронная сеть – это математическая модель, построенная по принципу сетей нервных клеток живого организма. Нейронная сеть не программируется в привычном понимании этого слова, ее необходимо обучать. Такая возможность – одно из главных преимуществ

над традиционными алгоритмами. В настоящее время нейронные сети широко используются в практических целях, например в задачах прогнозирования, распознавания образов, автоматизации процессов и др.

Алгоритм для обучения нейронной сети для решения поставленной задачи состоит из 4 этапов.

Этап 1. Извлечение признаков. Признаками можно считать некоторые особенности изображения, которые получаются из пикселей [1]. Пиксель – это наименьший логический элемент двумерного цифрового изображения в растровой графике [2]. Каждый пиксель представляет собой число или набор чисел, характеризующих глубину цвета, что позволяет нейронной сети выявлять некоторые закономерности.

Этап 2. Разметка изображения. Для того чтобы нейронная сеть смогла отличать объекты друг от друга, необходимо собрать множество изображений (датасет), в которых будут присутствовать требуемые объекты. После этого объекты на каждом изображении выделяются в ог-

раничивающую рамку (4 координаты, записанные в пикселях на изображении). Таким образом, получаются размеченные данные для дальнейшего обучения. Такой метод называется обучением с учителем [3].

Этап 3. Обучение модели. На вход сети поступает набор признаков, а на выходе получают метки соответствующих изображений. Цель обучения – научить сеть по поступающим признакам выделять необходимую метку, соответствующую изображению, с которого были получены признаки [4].

Этап 4. Определение классов на новых изображениях. После обучения сеть можно использовать для распознавания новых изображений.

На сегодняшний день одной из самых перспективных моделей для детекции средств индивидуальной защиты является архитектура нейронных сетей YOLO. Отличительной осо-

бенностью YOLO является высокая скорость обработки изображений [5]. Многие другие алгоритмы разбивают изображения на квадратные области, затем классифицируют области на наличие необходимого объекта и в конечном итоге классифицируют объект, из-за чего изображение «просматривается» дважды. В YOLO же используется сжатие изображения до матрицы 13×13 , где в каждой клетке записана информация о наличии объекта, что позволяет «просматривать» изображение один раз [6].

YOLOv5 – это пятая усовершенствованная версия YOLO. Также YOLOv5 имеет несколько вариантов реализации: YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x. Данные реализации отличаются количеством слоев нейронов (см. рис. 1), что позволяет выбирать между скоростью и точностью обучения (см. рис. 2) [7].

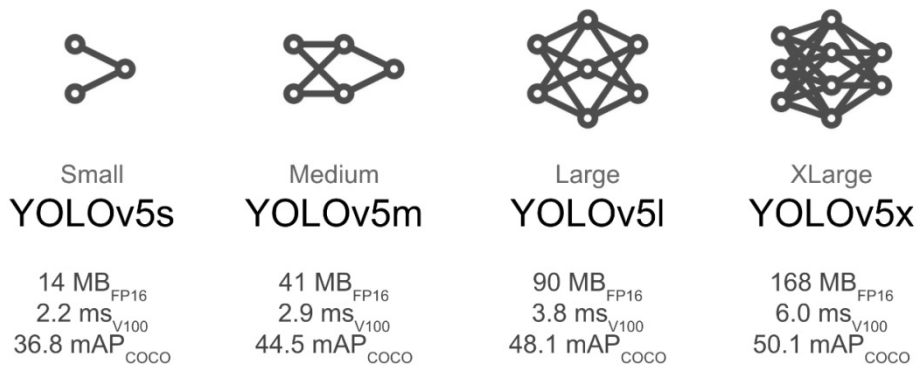


Рис. 1. Отличие реализаций нейронной сети YOLOv5
Fig. 1. The difference between the implementations YOLOv5

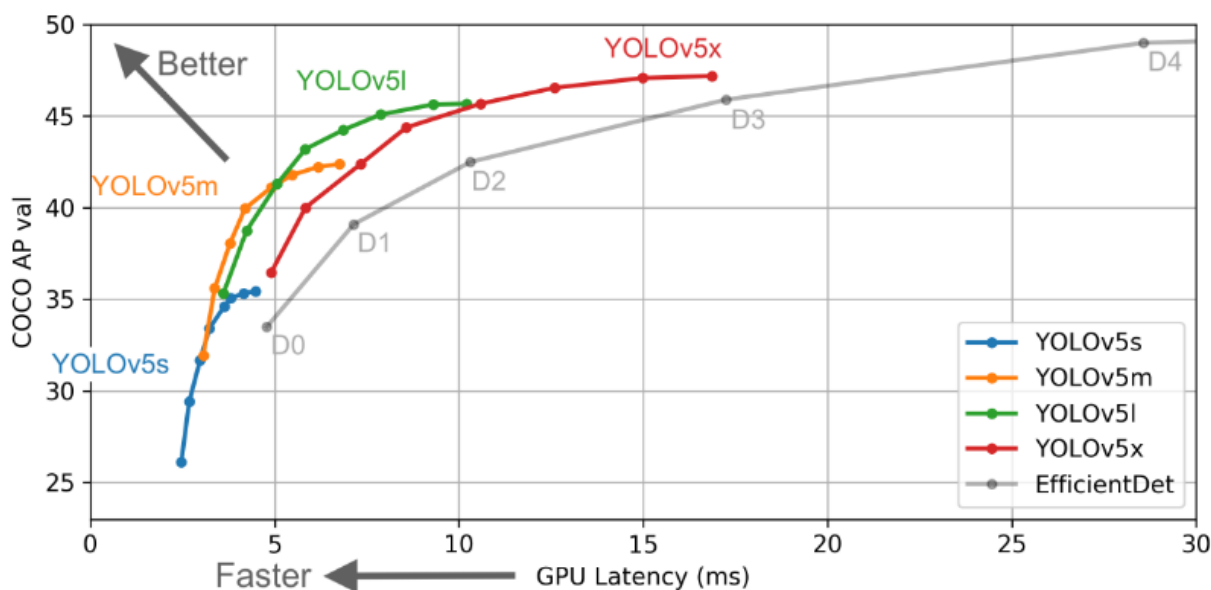


Рис. 2. Зависимость нейронных сетей от скорости обучения и точности определения
Fig. 2. Dependence of neural networks on the speed of definition and solution

Таким образом, одной из подзадач решения задач распознавания образов является подбор наиболее подходящей архитектуры сети для решения конкретной задачи [8].

Формирование датасета и настройка параметров YOLOv5

При обучении нейронных сетей требуются большие вычислительные мощности. Для этого часто используют вычислительные мощности видеокарт или специальных процессоров [9]. Компания Google предоставляет доступ к своим вычислительным мощностям, в том числе и для обучения нейронных сетей, посредством собственной разработки – блокнота Colab.

Для обучения YOLOv5 был составлен датасет из 2973 изображений и 6 классов, позволяющих характеризовать защищенность людей на улице в период нестабильной эпидемиологической обстановки и соблюдения техники безопасности рабочими на опасных объектах:

- с маской;
- без маски;
- маска надета неправильно;
- с головным убором;
- без головного убора;
- спец. одежда.

YOLOv5 позволяет выбрать несколько параметров при обучении [10], например:

- `img` – размер изображения, подаваемого на вход, т. е. исходное изображение будет преобразовано к необходимому;

- `batch` – количество изображений, подаваемых на вход;

- `epoch` – количество эпох для обучения.

Для обучения нейронной сети разработчикам YOLO рекомендуется использовать следующие параметры:

- больше 1500 изображений на каждый класс;

- больше 10000 представлений (помеченных объектов) каждого класса;

- разнообразие изображений (время суток, разные сезоны года, условия погоды, разного рода освещение, разные углы и т. д.);

- согласованность разметки (все экземпляры классов на всех изображениях должны быть отмечены);

- точность разметки (рамка должна вплотную прилегать к контурам объекта);

- добавить фоновые изображения (изображения без объектов, необходимо для уменьшения числа ложных срабатываний) [11].

Таким образом, меняя настройки нейронной сети и улучшая датасет, можно добиться наилучших результатов в распознавании.

На тестовом датасете будет обучено несколько вариантов нейронной сети со следующими настройками:

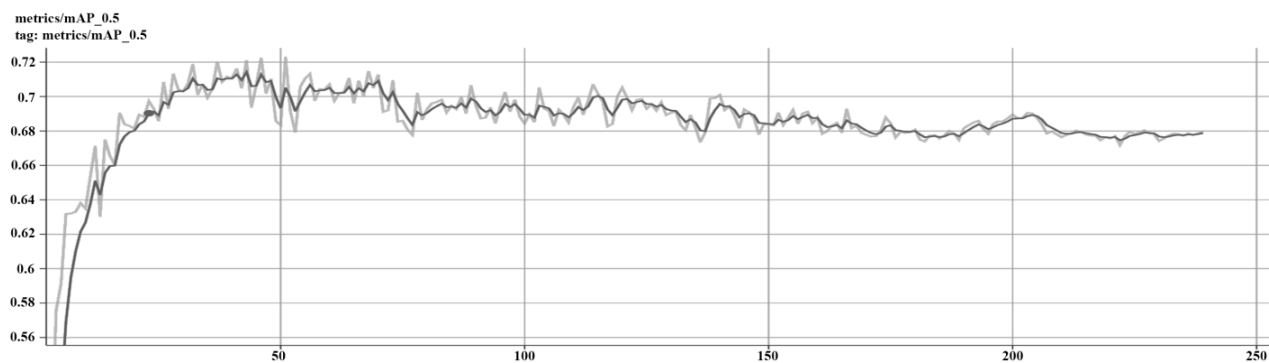
- `YOLOv5s --img 416 --batch 16 --epochs 100;`
- `YOLOv5s --img 416 --batch 16 --epochs 300;`
- `YOLOv5s --img 416 --batch 16 --epochs 40;`
- `YOLOv5l --img 416 --batch 16 --epochs 50.`

Обучение нейронной сети и анализ результатов

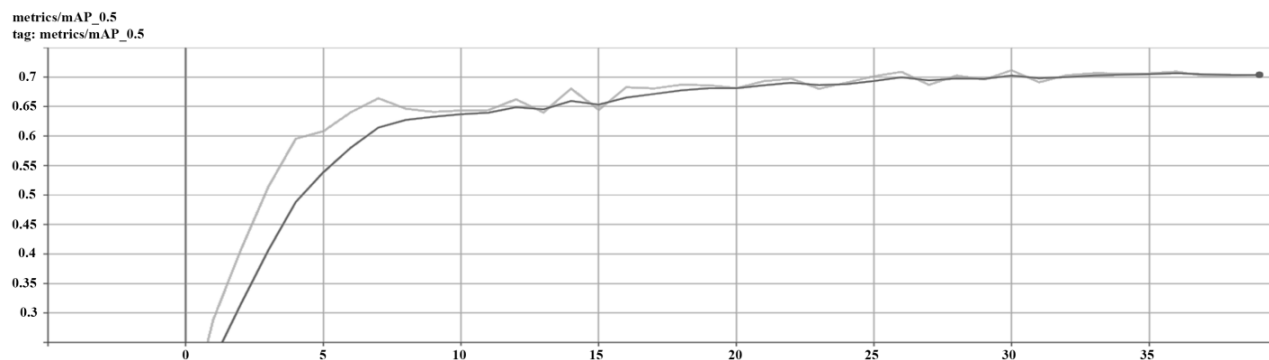
Скорость обучения нейронной сети зависит от количества изображений, классов и выбранных настроек [12]. Время, затраченное на обучение, может варьироваться от нескольких минут до нескольких дней. Для оптимизации скорости обучения, как один из вариантов, измеряют точность определения от количества эпох. Не всегда большое количество эпох позволяет добиться лучшего результата. Часто точность определения может падать из-за переобучения нейронной сети [13].

Одной из основных метрик эффективности нейронной сети является определение средней точности (Mean Average Precision – mAP). Точность определяется следующим образом: на проверочной выборке (размеченные данные, которые нейронная сеть «не видела») сеть пытается распознать объект и наложить ограничивающую рамку, после этого происходит сравнение ограничивающих рамок, размеченных заранее, и рамок от нейронной сети [14]. Затем по перекрывающейся площади можно вычислить точность [15, 16].

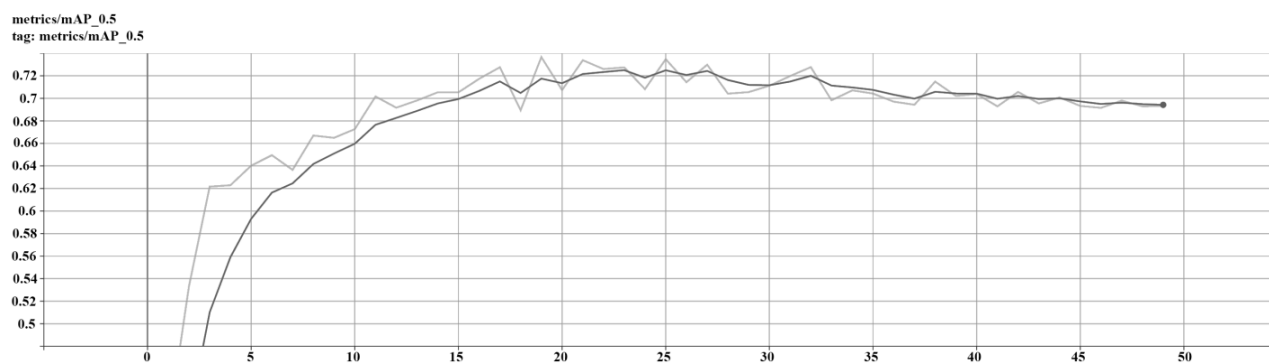
На рис. 3 представлены зависимости средней точности распознавания от номера эпохи для различных архитектур YOLOv5. На рис. 3, *a* и *b* видно, что точность начинает падать после первых 30–40 эпох. Данный процесс происходит, как правило, из-за переобученности сети. Также на рис. 3, *a* следует, что обучение нейронной сети было закончено на 240-й эпохе, хотя изначально было запущено 300 эпох. Такое произошло из-за того, что последние 100 эпох не приносили положительного результата и дальнейшее обучение сети бесполезно.



а



б



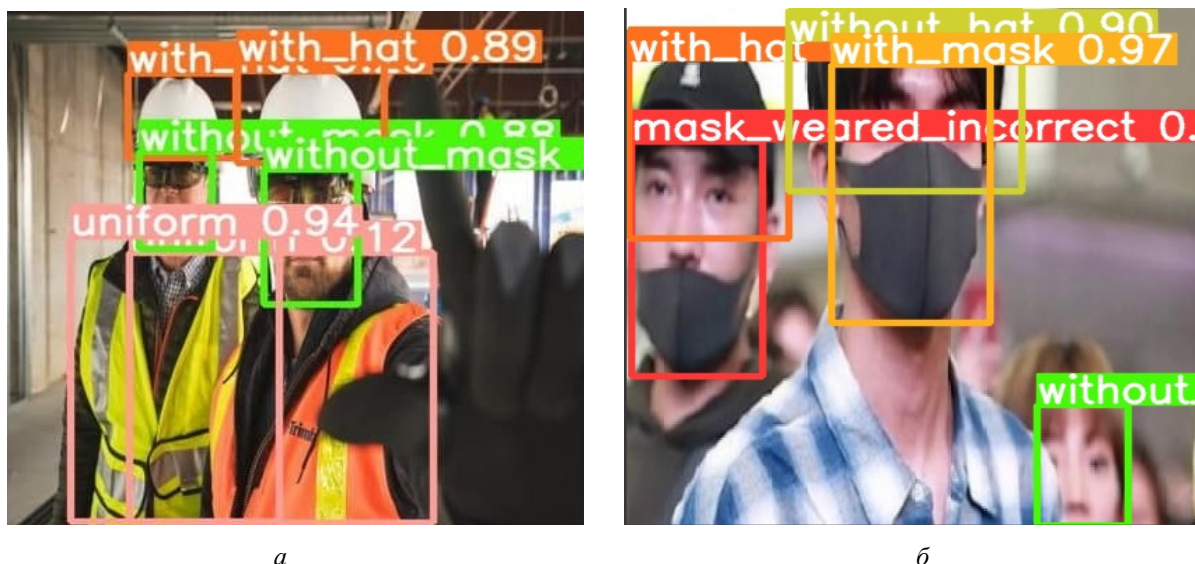
в

Рис. 3. Зависимость средней точности распознавания от номера эпохи для различных архитектур YOLOv5:
а – YOLOv5s 240 эпох, б – YOLOv5s 40 эпох, в – YOLOv5l 50 эпох

Fig. 3. Dependence of the average recognition accuracy on the epoch number for different architectures YOLOv5:
а – YOLOv5s 240 epochs, б – YOLOv5s 40 epochs, в – YOLOv5l 50 epochs

На рис. 4 представлены результаты распознавания наличия средств индивидуальной защиты на цифровых изображениях тестовой выборки. Из данных результатов следует, что разработанная интеллектуальная система с высокой точностью распознала у людей наличие

(with_hat) и отсутствие головного убора (without_hat), наличие спец. одежды (uniform), наличие (with_mask) и отсутствие масок (without_mask), а также неправильное ношение маски (mask_wearred_incorrect).



а

б

Рис. 4. Пример распознавания объектов на изображениях

Fig. 4. An example of object recognition in an image

Заключение

Проведенные исследования показывают высокую эффективность использования нейронной сети YOLOv5 для распознавания наличия средств индивидуальной защиты. В среднем точность распознавания составляет 0,7, что неплохо для данного класса задач. Как правило, низкая точность распознавания изображений проявляется на фотоизображениях достаточно низкого качества, а также где средство индивидуальной защиты плохо видно.

Для повышения точности распознавания необходимо продолжить дальнейшую работу по повышению качества датасета, а именно увеличения общего количества изображений и качества изображений на каждый класс детекции.

В ближайшей перспективе планируется распознавание новых классов детекции объектов индивидуальной защиты, в частности одежды, обуви для различных профессий, защитных перчаток. Одной из перспективных задач также является подтверждение факта прохождения процедуры дезинфекции и соблюдения других регламентных процедур персоналом предприятия на основе цифрового видеоряда.

Библиографические ссылки

1. Шмаглит Л. А., Голубев М. Н. Использование информации о цвете в алгоритме выделения лиц на изображениях // Материалы XX Международной конференции по компьютерной графике и зрению «ГрафиКон'2010». 2010. С. 331–332.
2. Визильтер Ю. В. Обработка и анализ изображений в задачах машинного зрения. М. : ФИЗМАТКН, 2010. 245 с.
3. Lin T.-Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollar P. and Zitnick C. L. Microsoft coco: Common objects in context // In European Conference on Computer Vision. 2014. Pp. 740-755.
4. Филичкин С. А. Применение нейронных сетей для распознавания поведенческих функций человека // Экология и безопасность жизнедеятельности : сборник статей XXI Международной научно-практической конференции. Пенза: Пензенский государственный аграрный университет, 2021. С. 206–210.
5. Elian F., Hariadi F. I. and Arsyad M. I. Implementation of computer vision algorithms for position correction of chip-mounter machine // 2017 International Symposium on Electronics and Smart Devices (ISESD). 2017. Pp. 90-94. DOI 10.1109/ISESD.2017.8253311.
6. Redmon J., Divvala S., Girshick R. and Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, pp. 779-788. DOI: 10.1109/CVPR.2016.91.
7. Официальный репозиторий YOLOv5. URL: <https://github.com/ultralytics/yolov5> (дата обращения: 26.02.2022).
8. Ворожцова Н. А., Вологдин С. В. Подготовка набора данных для распознавания показаний с фотографий лицевых панелей приборов учета электроэнергии // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2020. № 4. С. 121–126. DOI 10.25586/RNU.V9187.20.04. P.121.
9. Vorozhtsova N., Shushkov I., Vologdin S. System approach to development of Intellectual Information Mobile System for Electric Power Metering // E3S Web of Conferences, Irkutsk, 27–29 мая 2019 года. Irkutsk: EDP Sciences. 2019. 114. P. 01004. DOI 10.1051/e3sconf/201911401004.
10. Yaşar F. G., Kusetoçullari H. Underwater human body detection using computer vision algorithms //

2018 26th Signal Processing and Communications Applications Conference (SIU). 2018. DOI: 10.1109/SIU.2018.8404305.

11. Sanket J. Mankar, Manoj Demde, Prashant Sharma. Design of computer vision intelligent system for lane detection // 2016 Online International Conference on Green Engineering and Technologies (IC-GET). 2016. Pp. 1-3. DOI: 10.1109/GET.2016.7916843.

12. Боков П. А., Кравченко П. Д. Экспериментальный анализ точности и производительности разновидностей архитектур YOLO для задач компьютерного зрения. М. : Техносфера, 2018. 47 с.

13. Senthilkumaran N., Rajesh R. Edge Detection Techniques for Image Segmentation // International Journal of Recent Trends in Engineering. 2009. Vol. 1, No. 2. Pp. 250-254.

14. Sowmya B., Sheelarani B. Color Image Segmentation Using Soft Computing Techniques // International Journal of Soft Computing Applications. 2009. Issue 4. pp. 69-80.

15. Umesh Sehgal Edge detection techniques in digital image processing using Fuzzy Logic // International Journal of Research in IT and Management. 2011. Vol.1, Issue 3. Pp. 61-66.

16. Yu, X, Bui, T.D. Robust Estimation for Range Image Segmentation and Reconstruction // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1994. Vol. 16, Issue 5. pp. 530-538.

References

1. Shmaglit L.A., Golubev M.N. *Ispol'zovanie informatsii o tsvete v algoritme vydeleniya lits na izobrazheniyakh* [Using color information in the algorithm for selecting faces in images]. *Materialy 20-j Mezhdunarodnoj Konferencii po Komp'yuternoj Grafike i Zreniyu «Grafikon'2010»* [Proc. Materials of the XX Conference on Computer Graphics and Vision "Graphi-Con'2010"]. 2010, pp. 331-332 (in Russ.).

2. Vizilter Yu.V. *Obrabotka i analiz izobrazhenij v zadachah mashinnogo zreniya* [Image processing and analysis in machine vision problems]. Moscow, FIZMATKN, 2010, 245 p. (in Russ.).

3. Lin T.-Y., Maire P., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., and Zitnick C. L.. Microsoft coco: Common objects in context. In European Conference on Computer Vision, 2014, pp. 740-755.

4. Filichkin S.A. *Primenenie nejronnyj setej dlya raspoznavaniya povedencheskih funkcyj cheloveka* [Application of neural networks for recognition of human behavioral functions]. *Ekologiya i bezopasnost' zhiznedeyatel'nosti: Sbornik statej XXI Mezhdunarodnoj nauchno-prakticheskoy konferencii. Penza: Penzenskij gosudarstvennyj agrarnyj universitet* [Proc. Ecology and life safety : collection of articles of the XXI International Scientific and Practical Conference], 2021, pp. 206-210 (in Russ.).

5. Elian F., Hariadi F.I. and Arsyad M.I. Implementation of computer vision algorithms for position correction of chip-mounter machine. 2017 International Sym-

posium on Electronics and Smart Devices (ISESD). 2017. Pp. 90-94. DOI 10.1109/ISESD.2017.8253311.

6. Redmon J., Divvala S., Girshick R. and Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, pp. 779-788. DOI: 10.1109/CVPR.2016.91.

7. *Ofitsial'nyi repozitorii YOLOv5* [Official YOLOv5 repository] (in Russ.). Available at: <https://github.com/ultralytics/yolov5> (accessed 26.02.2022).

8. Vorozhtsova N.A., Vologdin S.V. *Podgotovka nabora dannykh dlya raspoznavaniya pokazanii s fotografiy litsevykh panelei priborov ucheta elektroenergii* [Preparation of a data set for recognition of indications from photographs of the front panels of electricity metering devices]. *Vestnik Rossiiskogo novogo universiteta. Seriya: Slozhnye sistemy: modeli, analiz i upravlenie*, 2020, no. 4, pp. 121-126 (in Russ.). DOI 10.25586/RNU.V9187.20.04. P.121.

9. Vorozhtsova N., Shushkov I., Vologdin S. System approach to development of Intellectual Information Mobile System for Electric Power Metering. In E3S Web of Conferences, Irkutsk, May 27-29, 2019, Irkutsk: EDP Sciences, 2019, 01004 p. DOI 10.1051/e3sconf/201911401004.

10. Fatma Günseli Yaşar, Hüseyin Kusetoğullari. Underwater human body detection using computer vision algorithms. In IEEE, 2018 26th Signal Processing and Communications Applications Conference (SIU), 09 July 2018, DOI: 10.1109/SIU.2018.8404305.

11. Sanket J. Mankar, Manoj Demde, Prashant Sharma. Design of computer vision intelligent system for lane detection. IEEE, 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 04 May 2017, DOI: 10.1109/GET .2016.7916843.

12. Bokov P.A., Kravchenya P.D. *Eksperimental'nyi analiz tochnosti i proizvoditel'nosti raznovidnostei arkhitektur YOLO dlya zadach komp'yuternogo zreniya* [Experimental analysis of the accuracy and performance of various YOLO architectures for computer vision tasks]. Moscow, Technosfera, 2018. 47 p. (in Russ.).

13. Senthilkumaran N., Rajesh R. Edge Detection Techniques for Image Segmentation. In A Survey of Soft Computing Approaches, International Journal of Recent Trends in Engineering, 2009, vol. 1, no. 2, pp. 250-254.

14. Sowmya B., Sheelarani B. Color Image Segmentation Using Soft Computing Techniques. In International Journal of Soft Computing Applications, 2009, Issue 4, pp. 69–80.

15. Umesh Sehgal Edge detection techniques in digital image processing using Fuzzy Logic. In International Journal of Research in IT and Management, 2011, vol.1, Issue 3, pp. 61-66.

16. Yu, X, Bui, T.D. Robust Estimation for Range Image Segmentation and Reconstruction. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, vol. 16, Issue 5 pp. 530-538.

Using the YOLOv5 Neural Network to Recognize the Presence of Personal Protective Equipment

S. A. Filichkin, Master Degree student, Kalashnikov ISTU, Izhevsk, Russia

S. V. Vologdin, DSc in Engineering, Associate Professor, Kalashnikov ISTU, Izhevsk, Russia

The article discusses the application of the YOLO neural network for the automatic detection and identification of personal protective equipment for enterprise personnel in digital images. A comparative analysis of various YOLOv5 network architectures is given, depending on the neural network learning rate and recognition accuracy. The tools for training a neural network in Python within the Google Colab environment, which allows efficient use of remote computing power of GPUs, are considered.

An example of network training based on collected dataset with six classes for recognizing the presence of personal protective equipment (with a mask; without a mask; the mask is worn incorrectly; with a head protector; without a head protector; special clothing) is considered. The settings of the YOLOv5 neural network and recommendations for improving object recognition by supplementing the dataset with certain condition images are described. Four variants of network training and different number of epochs were produced using the YOLOv5s and YOLOv5l architecture. The optimal neural network architecture and the number of required epochs has been selected. The concept of average accuracy (Mean Average Precision – mAP) is considered and a comparative analysis of the prediction accuracy from the selected network architecture and settings is given.

A practical example of recognizing the presence of personal protective equipment in digital images using a trained YOLOv5 neural network is given. On average, the recognition accuracy is 0.7, which is not a bad result for solving this class of problems. The paper deals with issues of improving the accuracy of object recognition. Plans for the detection of previously detected and new classes of personal protective equipment are given, as well as confirmation of the fact that the personnel of enterprises have passed the regulatory procedures based on digital video sequence.

Keywords: Computer vision, artificial intelligence, pattern recognition, neural network, image markup, YOLO.

Получено: 28.03.22