

10.37882/2223–2966.2021.04–2.12

ТРАНСФЕРНЫЙ ПОДХОД К ОБУЧЕНИЮ СВЁРТОЧНЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ ДИАГНОСТИКИ ПУЛЬМОНОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

TRANSFER APPROACH TO THE TRAINING OF CONVOLUTIONAL ARTIFICIAL NEURAL NETWORKS IN THE TASK OF DIAGNOSING PULMONOLOGICAL DISEASES

**V. Maslennikov
S. Daeva**

Summary. The article is devoted to the study of the application of the transfer approach to the training of convolutional artificial neural networks when working with medical graphic materials. An algorithm for developing a classifier based on convolutional neural networks for detecting pneumonia on digital images of chest X-rays is proposed. The study shows the advantages of the considered approach to the training of convolutional neural networks, as well as the advantages of the classifier developed on its basis over analogues.

Keywords: artificial intelligence, transfer learning, convolutional neural networks, pneumonia detection, classifier.

Масленников Владимир Владимирович
Российский технологический университет МИРЭА
vldmsn@yahoo.com

Даева Софья Георгиевна
К. ф. - м. н., доцент, Российский технологический университет МИРЭА

Аннотация. Статья посвящена исследованию вопросов применения трансферного подхода к обучению свёрточных искусственных нейронных сетей при работе с медицинскими графическими материалами. Предложен алгоритм разработки классификатора на основе свёрточных нейронных сетей для обнаружения пневмонии на цифровых изображениях рентгеновских снимков грудных клеток. Проведённое исследование показывает преимущества рассматриваемого подхода к обучению свёрточных нейронных сетей, а также достоинства разработанного на его основе классификатора перед аналогами.

Ключевые слова: искусственный интеллект, трансферное обучение, свёрточные нейронные сети, обнаружение пневмонии, классификатор.

Введение

С каждым годом всемирное развитие медицины в области производства высокотехнологичных и инновационных продуктов обеспечивает общество всё более качественными и доступными вариантами поддержки и укрепления здоровья населения нашей планеты. Несмотря на это, не все из них могут гарантировать высокоточные методы определения у человека разного рода заболеваний. Следует отметить и влияние так называемого человеческого фактора, ведь не всегда врач в силу каких-либо причин и особенностей способен увидеть в живом организме развивающуюся болезнь. Так, например, воспаление лёгочных тканей или пневмония достаточно часто имеет свойство носить скрытый характер, что может приводить к смертельному исходу в случае её позднего обнаружения [1]. В связи с этим для разрешения подобных проблем применяются разнообразные технологии, работающие на базе искусственных нейронных сетей [2, 3]. Одной из таких технологий является компьютерное зрение, используемое для создания машин, способных обнаруживать, отслеживать и классифици-

ровать различные типы объектов. Если задачи обнаружения и отслеживания являются основополагающими в рассматриваемой технологии, то задача классификации есть второстепенная и представляет собой соотношение исследуемого объекта, как вектора значений признаков, с определённым классом. Однако, решение задач классификации не всегда даёт желаемые результаты, поскольку очень часто возникает ситуация, при которой количество значений признаков оказывается недостаточным для точной классификации, что, в свою очередь, может приводить к «семантическому разрыву», то есть к отсутствию понимания информации, заключённой в данных, и, соответственно, к возможному соотношению объекта с несколькими классами. В итоге эта ошибка ведёт к неудовлетворительно низкой точности работы нейронной сети и, как следствие, приводит к неверному диагнозу у пациента. Во избежание подобных проблем в последние годы при разработке классификаторов активно практикуется использование трансферного обучения, находящее своё применение не только в медицинской области, но и во многих других сферах, начиная от обычных прикладных задач, и заканчивая задачами в нейролингвистическом про-

граммировании, например, задачами предсказания последующих слов в предложении [4].

В предлагаемой автором статье исследуются аспекты использования трансферного обучения при работе с медицинскими графическими материалами, в частности при выявлении пневмонии на рентгеновских снимках грудных клеток, а также рассматривается его влияние на точностные характеристики искусственных нейронных сетей.

1. Теоретическая часть

1.1. Нейронные сети и виды их архитектур

В современном мире искусственные нейронные сети — это быстро развивающийся механизм, который используется для обучения компьютеров, роботов и аналитических систем разумному человеческому мышлению [5] за счёт нейросетевых методов, основанных на применении различных типов нейронных сетей. Так как многие другие изобретения человека в научно-технической сфере исходят от идей по жизнедеятельности биологических организмов, то и любая искусственная нейронная сеть, в частности её архитектура, берёт своё начало от естественных «живых» нейронных сетей. По этой причине они имеют несколько схожий принцип действия.

Приводя в пример схему работы нейронной сети, стоит отметить, что первоначально на входной слой сети поступает определённый набор информации в виде сигналов, передающихся при помощи синапсов с собственными весовыми коэффициентами следующему слою нейронов. При этом информация, полученная каждым из нейронов — это сумма всех данных, перемноженных на свой весовой коэффициент. Затем итоговое значение подставляется в функцию активации, что и формирует выходную информацию, следующую до конечного выхода. Это есть последний этап в алгоритме действия сети.

Несмотря на то, что представленная схема является общей для всех нейронных сетей, их архитектура может быть совершенно разной. Вследствие этого стоит кратко рассмотреть некоторые виды таких архитектур.

1. Перцептрон, имеющий самую простую архитектуру и представляющий собой математическую или компьютерную модель восприятия информации мозгом человека [6]. Перцептрон состоит из сенсоров, ассоциативных элементов и одного сумматора. Процесс обучения такой архитектуры заключается в обычном изменении весовых коэффициентов, после чего перцептрон может

выполнять функции распознавания образов и обобщения понятий.

2. Нейронная сеть адаптивной резонансной теории, действующая по принципу соотнесения входного изображения с одним из классов в том случае, если оно достаточно соответствует прототипу данного класса. Такой эффект достигается благодаря базовой архитектуре сети, включающей группу входных обрабатывающих нейронов, группу распознающих нейронов и группу управляющих нейронов.
3. Нейронная сеть прямого распространения сигнала или сеть радиально-базисных функций, ключевой особенностью архитектуры которой является содержание в сети скрытого слоя радиально симметричных нейронов, способных преобразовывать расстояние от входного вектора до определённого «центра» по какому-либо нелинейному закону. На выходе такой сети образуется линейная комбинация нейронных параметров и радиально-базисных функций, находящих своё применение не только в задачах классификации, но и в прогнозировании временных рядов [7].
4. Свёрточная нейронная сеть, являющаяся одной из нескольких форм многослойных нейронных сетей и имеющая довольно широкое применение в работе с графической и звуковой информацией [8]. Главная идея архитектуры такой сети заключается в использовании общих синаптических коэффициентов, чередовании свёрточных слоёв со слоями подвыборки, а также в обеспечении двумерной связности нейронов в локальных рамках, что оказывает благоприятный эффект не только на точность, но и на скорость классификации предоставляемых на вход сети цифровых изображений.

С учётом вышесказанного можно сделать вывод о целесообразности использования именно свёрточной нейронной сети при работе с медицинскими графическими материалами.

1.2. Модель Mask R-CNN

При работе с цифровыми изображениями в качестве основы для выделения их элементов, может использоваться модель Mask R-CNN, разработанная при помощи высокоуровневого языка программирования Python версии 3.0, открытой нейросетевой библиотеки Keras, нацеленной на работу с сетями глубокого обучения, и библиотеки машинного обучения TensorFlow [9]. Такая модель даёт возможность генерировать ограничительные рамки и маски сегментации для каждого из экземпляров объекта на целевом изображении. Модель

Таблица 1. Параметры этапов обучения нейронной сети

№ этапа	Скорость обучения	Число эпох	Эпоха начала обучения	Слой	Аугментация
1	0.012	2	1	верхний	нет
2	0.006	6	3	все	да
3	0.0012	16	7	все	да

является стандартной свёрточной нейронной сетью, осуществляющей работу в роли экстрактора признаков: например, если на вход подаётся изображение размером 1024x1024x3, то свёрточная сеть может превратить его в карту признаков размером 32x32x2048 [10], используемую в дальнейшем в виде входных данных для последующих слоёв. Кроме этого, рассматриваемая модель содержит модуль, называемый Region Proposal Network (RPN). Благодаря большому количеству полей с предопределёнными ограничительными рамками заданной ширины и высоты RPN проводит сканирование карты признаков на предмет присутствия нужного объекта. Затем модель принимает предложенную RPN область с объектом в качестве входных данных и выводит классификацию с ограничивающей рамкой. На заключительном этапе работы модели, исходя из выбранных модулем положительных регионов, в качестве выходных данных генерируется маска пикселей размером 28x28, охватывающая все пиксели, которые занимает объект на изображении.

1.3. Трансферный подход к обучению

Один из ключевых аспектов в разработке любой нейронной сети — это процесс её обучения и, как сказано ранее, наиболее актуальным является трансферный подход. Главная идея этого подхода состоит в «решении поставленной проблемы на основе «готовых данных», полученных в результате решения аналогичных задач» [11, с. 8]. Как известно, свёрточные нейронные сети состоят из нескольких слоёв: входного, выходного и множества уровней скрытых слоёв, отвечающих за выполнение операций для изменения входных данных в целях изучения специфических для них характеристик. Например, среди скрытых слоёв нижние могут распознавать края, детали и другие мельчайших элементов на изображении. Последние же слои определяют категорию изображения. Трансферное обучение даёт возможность переобучить последний слой нейронной сети на небольшом пользовательском наборе изображений, что значительно сокращает время, затрачиваемое на обучение модели в целом, и в дальнейшем использовать оставшуюся модель распознавания по второму разу без изменения структуры [12].

1.4. Разработка классификатора для обнаружения пневмонии на рентгеновских снимках

Для разработки классификатора, предназначенного для решения задачи выявления воспалительного процесса в лёгких (пневмонии) на рентгеновских снимках грудных клеток, предлагается использовать следующий алгоритм, реализующий учёт совокупности всех вышеназванных факторов.

1. Взять готовый экземпляр свёрточной нейронной сети Mask R-CNN[9]. Осуществить импорт соответствующих библиотек для настройки классов, конфигурирующих обучение сети.
2. Применить для реализации процесса обучения модели Mask R-CNN трансферный подход с использованием предварительно обученных весов COCO (Common Objects in Context) [13], представляющих собой связи между нейронами, «натренированные» при помощи модели Mask R-CNN на наборах данных COCO от компании Microsoft и сохранённые в файле с иерархическим форматом данных.h5.

Процесс обучения нейронной сети предлагается разделить на три этапа, каждый из которых имеет соответствующие ему параметры (таблица 1).

На первом этапе предлагается использовать высокий коэффициент скорости обучения для того, чтобы максимально быстро «натренировать» верхний слой нейронной сети. Обучение на данном этапе проводится в течение двух эпох (итераций работы нейронной сети, которые включают в себя предъявление всех элементов обучающей выборки и проверку качества обучения на тестовой выборке).

Второй этап предлагается осуществлять с пониженным в два раза коэффициентом скорости обучения. Значение числа эпох устанавливается равным шести. Обучение продолжается с третьей эпохи, а в самом процессе задействуются все слои нейронной сети. При этом выполняется аугментация, позволяющая увеличить объём обучающей выборки посредством модификации

	val_loss	val_rpn_class_loss	val_rpn_bbox_loss	val_mrcnn_class_loss	val_mrcnn_bbox_loss	val_mrcnn_mask_loss
1	1.620174	0.019989	0.499384	0.224194	0.483136	0.393469
2	1.551982	0.025763	0.486674	0.213167	0.459840	0.366533
3	1.466383	0.017147	0.451255	0.188478	0.424972	0.384519
4	1.445297	0.014090	0.393461	0.221944	0.432236	0.383553
5	1.479642	0.014302	0.458050	0.212715	0.413680	0.380881
6	1.423630	0.012183	0.442380	0.181016	0.413563	0.374472
7	1.273788	0.011006	0.395211	0.128471	0.371854	0.367231
8	1.333833	0.011734	0.463713	0.111168	0.378438	0.368766
9	1.303161	0.011025	0.409615	0.128954	0.382112	0.371440
10	1.275731	0.010077	0.376540	0.144807	0.374715	0.369577
11	1.329650	0.010255	0.436387	0.138612	0.375533	0.368848
12	1.295761	0.010512	0.410356	0.130317	0.371950	0.372612
13	1.287467	0.010315	0.402464	0.144363	0.364404	0.365905
14	1.330533	0.009559	0.433231	0.153592	0.366991	0.367145
15	1.311978	0.010093	0.417572	0.153361	0.363597	0.367341
16	1.291173	0.009408	0.384802	0.154610	0.372381	0.369957

Рис. 1. Лучшая эпоха обучения по показателю потерь при валидации

имеющихся в наборе данных (изображений). В частности, для модификации изображений применяются:

- ◆ аффинные преобразования (масштабирование, перемещение, поворот, сдвиг);
- ◆ изменение яркости и контраста;
- ◆ изменение резкости;
- ◆ добавление размытия.

На третьем этапе предлагается снизить скорость обучения сети в пять раз. Обучение продолжается с седьмой эпохи, а общее число эпох в алгоритме фиксируется на значении шестнадцать. В процессе обучения по-прежнему участвуют все слои нейронной сети и также выполняется аугментация.

3. Определить, какая из эпох обучения была лучшей, то есть имела наименьший коэффициент потерь при валидации, для гарантирования высокого уровня обработки сетью входных типовых данных в дальнейшем.
4. Установить лучшую эпоху в качестве рабочей. Выполнить настройку отображения результатов, получаемых вследствие анализа входных графических материалов.

5. Приступить к проверке рентгеновских снимков лёгких из набора тестовых DICOM-изображений на наличие локальных очагов воспаления, возникающих при наличии бактериальной или вирусной пневмонии.

2. Экспериментальная часть

Обучение модели Mask R-CNN на основе трансферного подхода и предварительно обученных весов COCO на каждой из эпох даёт различные результаты коэффициента потерь при валидации. Однако, как видно из рисунка 1, наименьший коэффициент получается на эпохе № 7.

После рентгенографического исследования лёгких снимки не только печатаются на специальной плёнке, но и сохраняются в виде электронных файлов с расширением.dcm формата DICOM, содержащих цифровые изображения, которые могут подаваться на вход нейронной сети. В качестве обучающего графического материала с такими изображениями был взят набор данных RSNA (Pneumonia Detection Challenge), предо-

```
ID пациента.....: 19d67232-2d92-4486-b9bb-f8355f5d2a74
Возраст.....: 41
Пол.....: M
Исследование.....: CR
Способ осмотра.....: CHEST
Положение осмотра.....: AP
Размер изображения.....: 1024 x 1024, 114698 bytes
Расстояние между пикселями..: [0.168, 0.168]
```

```
Локальный очаг №1: точность - 0.99, x - 204.0, y - 504.0, ширина - 232.0, высота - 216.0
Локальный очаг №2: точность - 0.97, x - 600.0, y - 304.0, ширина - 212.0, высота - 220.0
Локальный очаг №3: точность - 0.96, x - 288.0, y - 188.0, ширина - 164.0, высота - 168.0
```

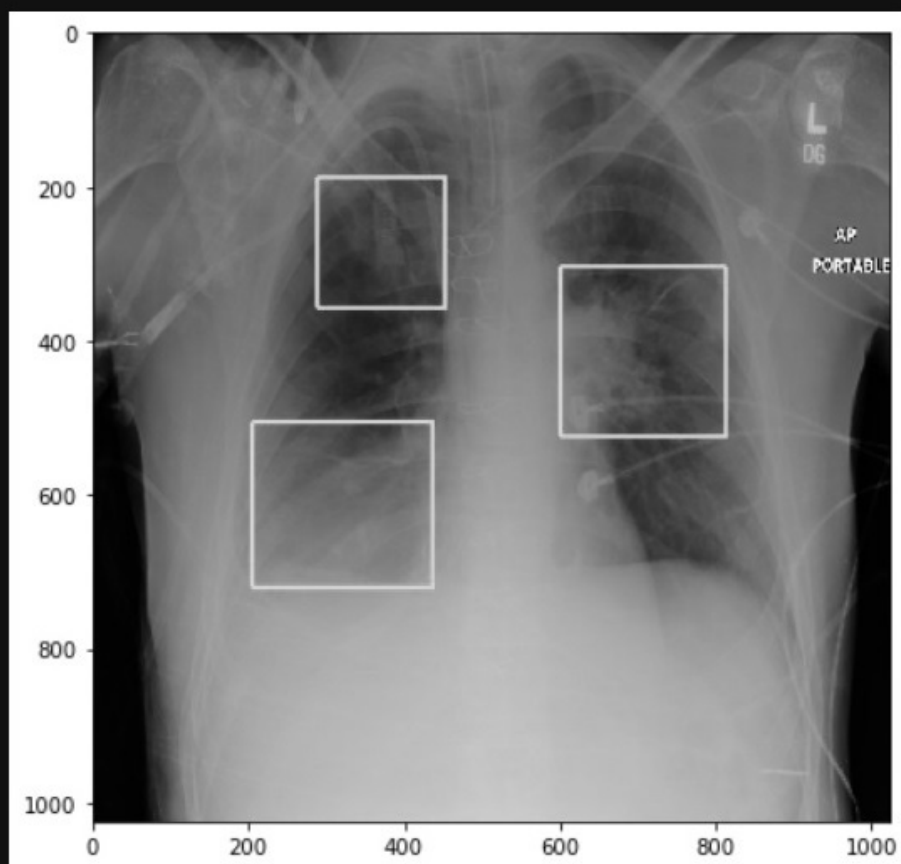


Рис. 2. Результат обработки изображения рентгеновского снимка грудной клетки

ставляемый системой Kaggle, которая занимается организацией конкурсов по исследованию информации [14]. Набор данных включает в себя 26684 обучающих и 3000 тестовых DICOM-изображений, а также *xlsx*-файл с предустановленными характеристиками для тренировочных изображений: идентификатором пациента, идентификатором наличия воспаления, координатами воспалительной области по осям x и y , её шириной и высотой.

Результаты обработки случайно взятого из набора изображения представлены на рисунке 2.

Как видно из рисунка 2, на нём указаны описание рентгеновского снимка, координаты и точность каждого найденного и классифицированного локального очага воспаления лёгких, а также сам снимок с обозначенными на нём областями очагов пневмонии. В зависимости от поступающего на вход сети изображения

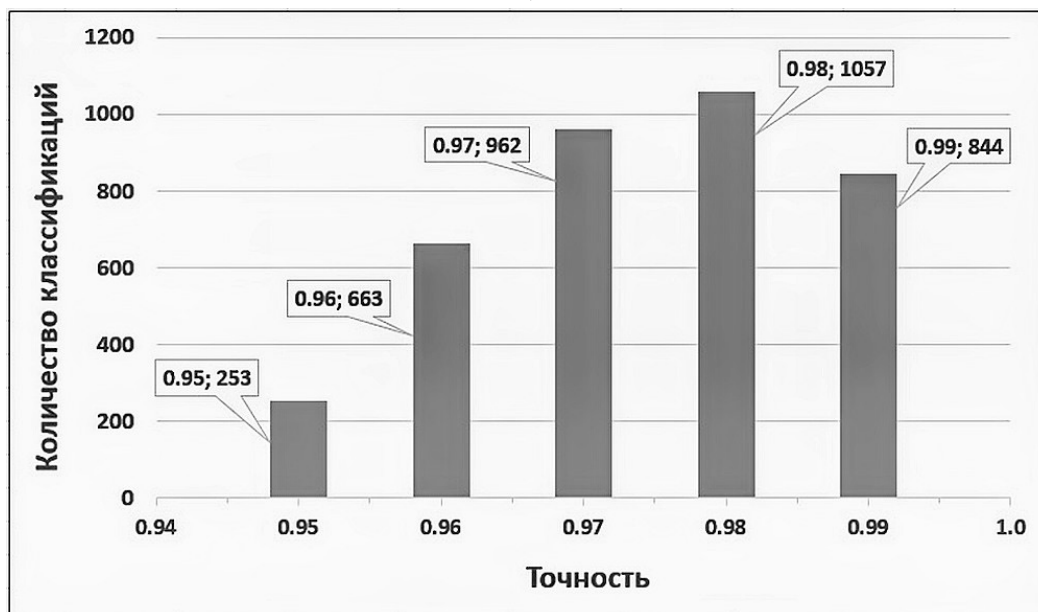


Рис. 3. Гистограмма общих результатов обработки изображений

лёгких таких выделяемых областей может быть большое количество. Однако при проверке нейронной сети на тестовом наборе изображений итоги классификации варьировались от одного до трёх очагов с минимальной точностью определения равной 95%. Результаты по обработке всех снимков из тестового набора подсчитаны в полной мере на следующей гистограмме (рисунок 3). На ней видно, что из всего набора с тремя тысячами тестовых изображений пневмония была классифицирована 253 раза с точностью 0.95 (95%), 663 раза с точностью 0.96 (96%) и 962 раза с точностью 0.97 (97%). Больше всего классификаций (1057 раз) проведено с точностью 0.98 (98%), а максимальная точность классификации была достигнута со значением 0.99 или 99% (844 раза).

В разных литературных источниках описывается большое количество нейросетевых средств, помогающих диагностировать пневмонию на рентгеновских снимках грудных клеток. Например, согласно [15] в одном из экспериментов рассматривается задача классификации воспаления лёгких, решаемая с использованием архитектуры Mask R-CNN и глубокого обучения на базе набора данных RSNA от Kaggle. Исследование

показало, что применение сочетания указанной сетевой модели и обучения типа «deep learning» даёт результат только в 78,06% точности, что на 16,94% меньше минимального показателя, полученного автором при помощи комбинации трансферного подхода с глубоким обучением. Данный факт показывает не только успешность такого подхода, но и целесообразность его эксплуатации в различных промышленных, и тем более, в медицинских системах, где крайне важную роль играют критерии скорости и точности решения различного рода насущных задач.

Заключение

В работе был предложен алгоритм разработки классификатора для обнаружения пневмонии лёгких на цифровых изображениях рентгеновских снимков грудных клеток, включающий в себя применение методики глубокого обучения с трансферным подходом. Результаты показали, что разработанный классификатор заметно превосходит аналоги по показателю точности. Кроме того, предлагаемый алгоритм является универсальным инструментом, что позволяет использовать его в абсолютно любом секторе мирового хозяйства.

ЛИТЕРАТУРА

1. Симптомы пневмонии скрытая форма // Здоровые лёгкие. [Электронный ресурс]: <https://pnevmonii.net.ru/simptomny-pnevmonii-skrytaya-forma.html> (дата обращения: 09.02.2021).
2. Wang Q, Bi S, Sun M, Wang Y, Wang D, Yang S (2019) Deep learning approach to peripheral leukocyte recognition. PLoS ONE14(6): e0218808. <https://doi.org/10.1371/journal.pone.0218808>.

3. Kadam, K., Ahirrao, S., Kaur, H., Phansalkar, S., & Pawar, A. (2019). Deep Learning Approach for Prediction of Pneumonia, *International Journal of Scientific & Technology Research*, 8 (10): 2986–2989.
4. The Illustrated BERT, ELMo and co. (How NLP Cracked Transfer Learning) // *GITHUB blog*. [Электронный ресурс]: <https://jalanmar.github.io/illustrated-bert> (дата обращения: 09.02.2021).
5. Щурина С.В., Данилов А.С. Искусственный интеллект как технологическая инновация для ускорения развития экономики. *Экономика. Налоги. Право*. 2019;12(3):125–133. DOI: 10.26794/1999–849X-2019–12–3–125–133.
6. Нейронные сети, перцептрон // *Университет ИТМО*. [Электронный ресурс]: https://neerc.ifmo.ru/wiki/index.php?title=Нейронные_сети,_перцептрон (дата обращения: 15.02.2021).
7. Сеть радиальных базисных функций // *Распознавание*. [Электронный ресурс]: http://www.machinelearning.ru/wiki/index.php?title=Сеть_радиальных_базисных_функций (дата обращения: 23.02.2021).
8. Гафаров Ф.М., Галимянов А.Ф. Искусственные нейронные сети и приложения. Казань: Казанский университет, 2018.
9. Mask R-CNN for Object Detection and Segmentation. [Электронный ресурс]: https://github.com/matterport/Mask_RCNN (дата обращения: 26.02.2021).
10. Buric, Matija & Pobar, Miran & Ivašić-Kos, Marina. (2018). Ball Detection Using Yolo and Mask R-CNN. P. 319–323. DOI: 10.1109/CSCI46756.2018.00068.
11. Казаков О.Д., Михеенко О.В. Трансфертное обучение и доменная адаптация на основе моделирования социально-экономических систем // *Бизнес-информатика*. 2020. Т. 14. № 2. С. 7–20. DOI: 10.17323/2587–814X.2020.2.7.20.
12. Трансферное обучение: Как классифицировать изображения с помощью машинного обучения Tensorflow // *Уроки по программированию, DevOps и другим IT-технологиям*. [Электронный ресурс]: <https://coderlessons.com/articles/programmirovanie/transfernoe-obuchenie-kak-klassifitsirovat-izobrazheniia-s-pomoshchiu-mashinnogo-obuchenii-tensorflow> (дата обращения: 02.03.2021).
13. Download COCO pre-trained weights. [Электронный ресурс]: https://github.com/matterport/Mask_RCNN/releases/download/v2.0/mask_rcnn_coco.h5 (дата обращения: 05.03.2021).
14. RSNA Pneumonia Detection Challenge // *Kaggle: Your Machine Learning and Data Science Community*. [Электронный ресурс]: <https://www.kaggle.com/rsna-pneumonia-detection-challenge> (дата обращения: 08.03.2021).
15. Elshennawy N.M., Ibrahim D.M. Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images. *Diagnostics* 2020, 10, 649. P. 2–3.

© Масленников Владимир Владимирович (vldmsn@yahoo.com), Даева Софья Георгиевна.
Журнал «Современная наука: актуальные проблемы теории и практики»



МИРЭА — Российский технологический университет