

## Алгоритм распознавания маркировки грузового контейнера с использованием глубоких нейронных сетей

*Р.А. Макаров*

*Казанский национальный исследовательский технический университет  
им. А.Н. Туполева-КАИ, Казань*

**Аннотация:** В данной статье реализован алгоритм распознавания маркировки грузового контейнера на основе глубоких нейронных сетей. Основным преимуществом предложенного алгоритма является отсутствие необходимости дополнительных преобразований изображения для локализации маркировки и сегментации символов. Для решения задачи локализации использовался алгоритм EAST. Для распознавания кода собственника и серийного номера контейнера были обучены модели нейронных сетей Faster R-CNN Resnet 50. В результате оценки точности работы алгоритма, были сделаны соответствующие выводы, а также сформулированы возможные варианты улучшения работы алгоритма.

**Ключевые слова:** машинное обучение, распознавание цифр, грузовые контейнеры, нейронная сеть, классификация изображений.

### Введение

На текущий момент реализация системы распознавания маркировки грузового контейнера включает в себя следующие этапы:

- 1) Локализация маркировки на изображении
- 2) Сегментация символов маркировки
- 3) Классификация символов

В результате этого удастся получить код собственника и серийный номер контейнера.

Исходя из анализа работ [1-3], видно, что для локализации и сегментации символов используются стандартные методы компьютерного зрения, а именно: бинаризация изображения, добавление размытия для получения более гладких границ символов, изменение яркости и контрастности.

Также исходя из анализа работ, видно, что наиболее слабым местом является алгоритм локализации и сегментации символов. Это связано с высокой вариативностью входных изображений для распознавания. В

---

результате различных вариантов установки камеры, а также различных погодных условий, яркость изображения может существенно изменяться. Также грузовые контейнеры не имеют стандартизированного цвета маркировки, в результате этого при локализации и сегментации на основе изменения яркости относительно фона могут возникнуть проблемы. Кроме того, грузовой контейнер подвержен воздействию внешних факторов, вследствие чего могут возникнуть дополнительные искажения маркировки грузового контейнера.

В результате возникает необходимость реализации алгоритма распознавания на основе глубоких нейронных сетей, для минимизации ошибки алгоритма классического компьютерного зрения с целью сегментации и локализации маркировки.

В данной статье рассматривается реализация алгоритма с использованием глубоких нейронных сетей. Основными целями работы являются проверка и оценка возможности использования только глубоких нейронных сетей для решения задачи распознавания маркировки грузового контейнера. Предложенный алгоритм позволяет снизить вероятность возникновения ошибки в результате этапа локализации и сегментации символов, что позволит достичь большей точности распознавания символов маркировки.

Также одной из особенностей предметной области является отсутствие наборов данных, содержащих грузовые контейнеры. Вследствие этого, для обучения и тестирования моделей необходимо собрать и разметить собственный набор данных.

### **Алгоритм локализации текста на изображении**

Для локализации текста на изображении нужно использовать реализацию алгоритма EAST (Efficient and Accurate Scene Text) [4]. Ключевым компонентом алгоритма является сверточная нейронная сеть,

которая обучена для прогнозирования нахождения элементов с текстом на изображении.

В данном случае, основной целью реализации алгоритма автор ставит возможность получения области с текстом без дополнительных преобразований исходного изображения. Это позволяет получить результат без необходимости изменения яркости изображения.

В исходной реализации, результатом работы алгоритма являются следующие данные:

- 1) Координаты ограничивающего прямоугольника с текстом
- 2) Точность определения области с текстом
- 3) Угол поворота текста.

Архитектура нейронной сети EAST представлена на рисунке 1.

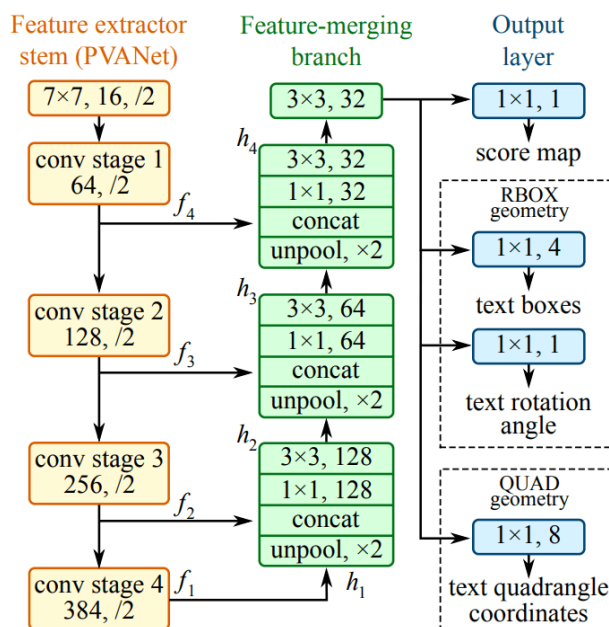


Рис. 1. – Архитектура нейронной сети EAST [4]

В данном случае для использования алгоритма было принято решение не использовать данные для поворота текста. Это обусловлено тем, что маркировка грузового контейнера не имеет поворота относительно контейнера и в данном случае это позволит упростить локализацию маркировки грузового контейнера.

Результат локализации текста на изображении контейнера представлен на рисунке 2. Исходя из него, видно, что модель успешно справляется с локализацией маркировки грузового контейнера на изображении.



Рис. 2. – Пример результата работы алгоритма EAST

### Распознавание серийного номера грузового контейнера

Для реализации распознавания серийного номера грузового контейнера использовалась модель, полученная в результате работы [5]. Используемая автором модель Faster R-CNN Resnet 50 реализует алгоритмы сегментации и распознавания. В результате этого отсутствует необходимость реализовывать алгоритм сегментации на основе преобразования изображения.

Пример результата работы модели представлен на рисунке 3.

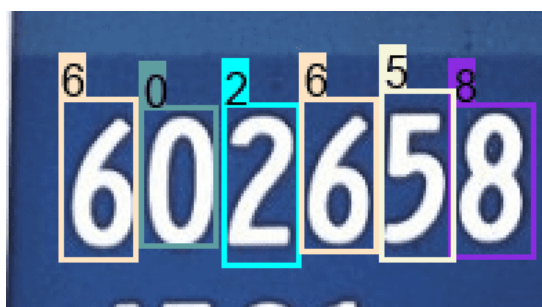


Рис. 3. – Пример результата распознавания серийного номера контейнера

Модель обучена на открытом наборе данных Street View House Numbers. Это связано с отсутствием данных с грузовыми контейнерами.

Результатом работы алгоритма являются следующие данные:

- 1) Координаты ограничивающего прямоугольника с цифрой
- 2) Точность распознавания цифры
- 3) Класс распознанной цифры

### **Распознавание кода собственника грузового контейнера**

Для реализации модели распознавания кода собственника также использовался алгоритм Faster R-CNN, данный алгоритм имеет сходство с алгоритмом RetinaNet [6]. Для обучения модели использовался открытый набор данных ПИТ 5К. Выбранный набор содержит 5000 изображений слов.

К преимуществам выбранного набора данных можно отнести высокую вариативность цветов с текстом и фоном, а также использование различных шрифтов. Это позволяет увеличить точность работы модели [7].

Пример данных для обучения представлен на рисунке 4.



Рис. 4. – Пример данных для обучения модели распознавания букв

Обучение модели производилось с использованием языка программирования Python и фреймворка Tensorflow. В данном случае использование фреймворка позволяет производить процесс обучения с использованием графического процессора, что увеличивает скорость обучения [8]. График функции потерь в процесс обучения представлен на рисунке 5. Исходя из него, видно, что модель обучалась около 39 тыс. итераций, среднее значение функции потерь не превышало 10%. В результате этого, можно сделать вывод, что модель успешно обучена.

---

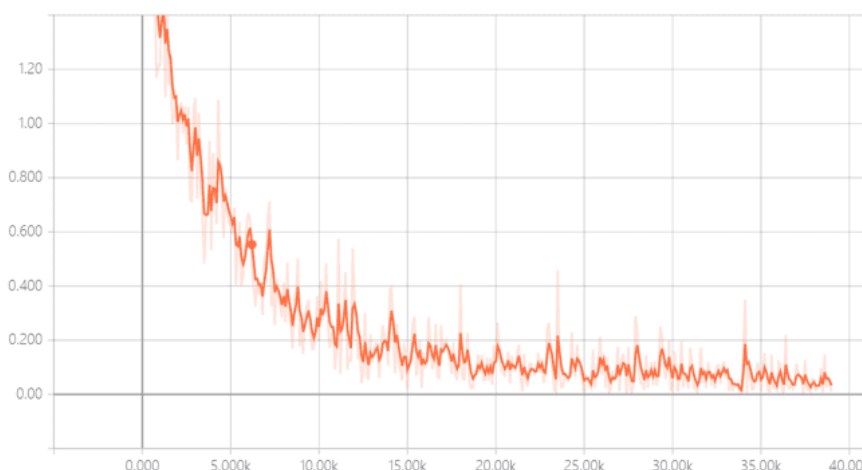


Рис. 5. – График функции потерь в процессе обучения модели

Пример работы модели по распознаванию кода собственника представлен на рисунке 6.

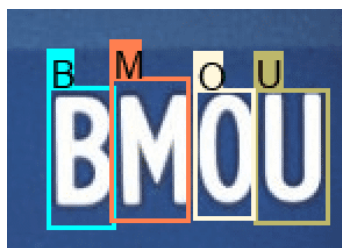


Рис. 6. – Пример работы модели по распознаванию букв

Исходя из рисунка 6, видно, что модель успешно распознает код собственника грузового контейнера.

### **Реализация алгоритма распознавания маркировки**

На рисунке 7 представлена схема предложенного алгоритма распознавания маркировки грузового контейнера. Исходя из схемы, видно, что первым этапом работы является локализация текста на изображении. Критерием успешной локализации является определение двух областей с текстом на одном уровне, также области должны располагаться выше, относительно других найденных областей. Данный критерий обусловлен тем, что маркировка грузового контейнера располагается в правом верхнем углу. После успешной локализации областей, распознавание серийного номера и кода собственника контейнера производится в двух параллельных потоках.



Это обусловлено тем, что в результате локализации, имеются две отдельных области с кодом собственника и серийным номером контейнера, реализация параллельной обработки фрагментов позволяет существенно сократить время работы алгоритма [9].

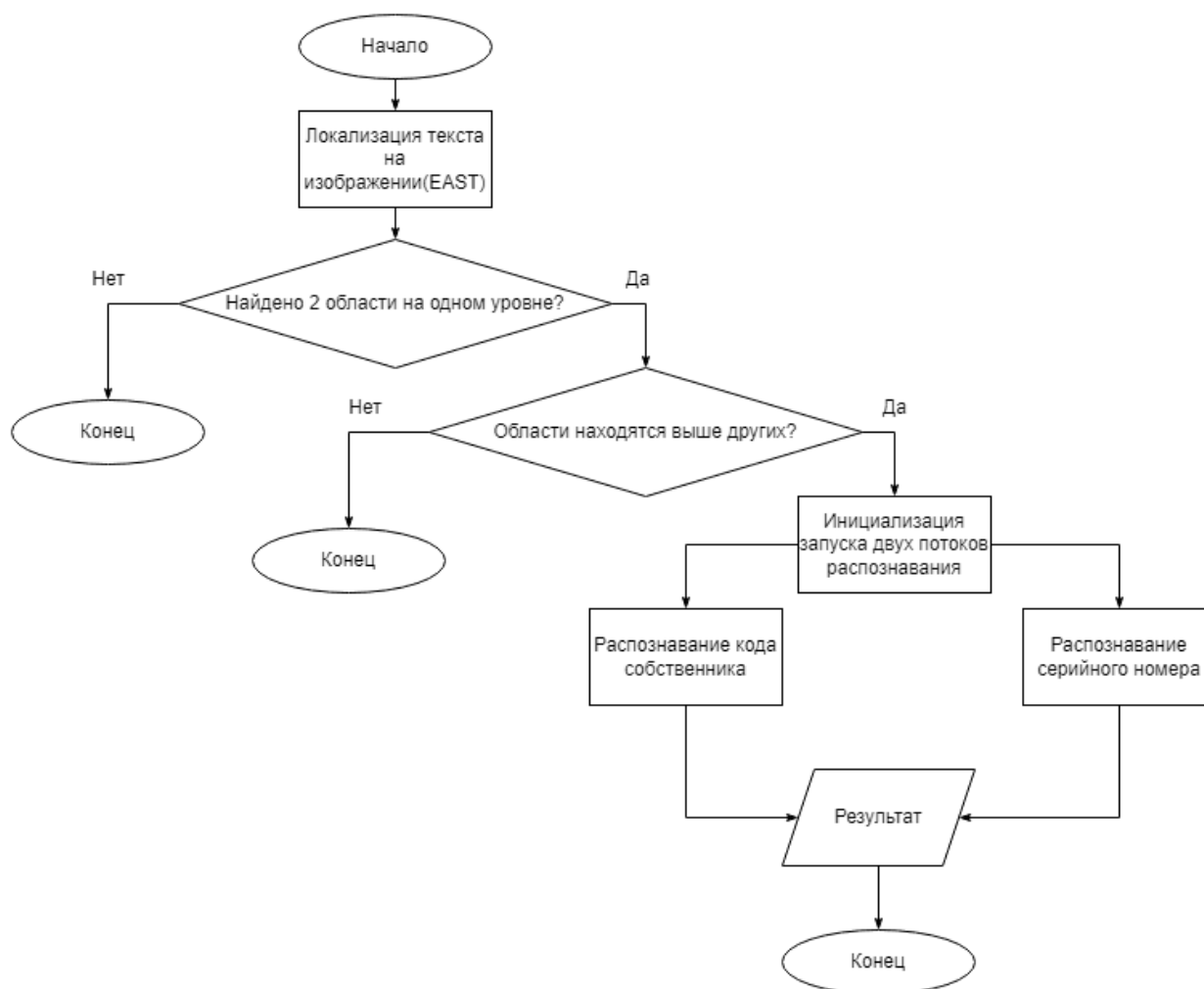


Рис. 7. – Схема работы алгоритма распознавания маркировки

Для программной реализации алгоритма распознавания использовался язык программирования Python и фреймворк Tensorflow. Обработка и распознавание производилось с использованием графического процессора. Для реализации распознавания кода собственника и серийного номера были реализованы две отдельные функции, вызов которых производился одновременно.

Результат работы алгоритма представлен на рисунке 8.

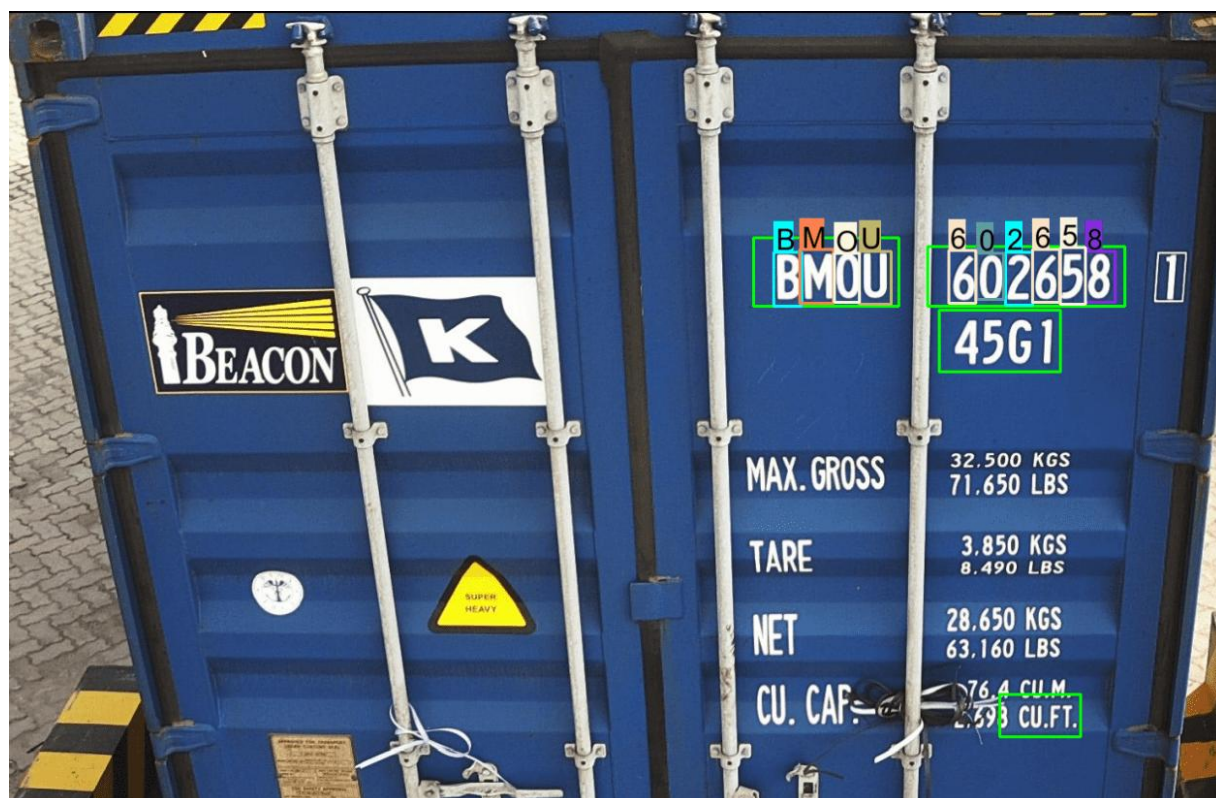


Рис. 8. – Результат работы алгоритма распознавания маркировки грузового контейнера

### Оценка точности работы алгоритма

Для оценки точности работы алгоритма, было собрано 115 изображений с грузовыми контейнерами, снятыми под прямым углом к камере. С целью корректной оценки, необходимо также определить точность работы алгоритма в случае поворота контейнера на изображении [10]. Для этого, на основе собранных данных, были сформированы дополнительные наборы с поворотами контейнеров на изображении и были выбраны следующие углы поворотов:

- 0 градусов
- 5-15 градусов
- 15-25 градусов
- 25-35 градусов
- 35-50 градусов



В данном случае, поворот осуществлялся на основе преобразования изображения. Шаг изменения обусловлен возможными вариантами установки камеры для реализации системы распознавания. Угол 50 градусов является максимальным критическим, в противном случае необходимо рассмотреть другое место для установки камеры.

Результат оценки точности работы алгоритма представлен в таблице 1. Исходя из нее, видно, что наибольшая точность распознавания обеспечивается в случае отсутствия поворота контейнера, с увеличением угла поворота точность распознавания понижается.

Таблица № 1

Точность работы алгоритма распознавания в зависимости от угла поворота грузового контейнера

Угол поворота	0 градусов	5-15 градусов	15-25 градусов	25-35 градусов	35-50 градусов
Точность распознавания	96,73%	94,56%	92,38%	90,21%	88,5%

### Заключение

В результате данной работы был реализован алгоритм распознавания маркировки грузового контейнера. Для реализации алгоритма были обучены модели для распознавания цифр и букв на изображении.

Главной целью работы – являлось разработка алгоритма распознавания с использованием глубоких нейронных сетей, что позволит не использовать алгоритмы компьютерного зрения для сегментации символов на изображении. В результате оценки точности работы алгоритма, можно сделать вывод, что алгоритм является работоспособным.

Для улучшения точности работы необходимо расширение обучающей выборки, а именно: увеличение обучающей выборки цифр и букв, путем добавления дополнительных данных с увеличенным углом поворота символа

и использование реальных данных для обучения моделей. Также, для улучшения работы алгоритма, можно рассмотреть применение весовых моделей изображения [11-13] с целью распознавания границ образов и сжатия данных.

В результате увеличения обучающей выборки, удастся повысить точность работы алгоритма при поворотах контейнера на изображении. Использование реальных данных позволит дополнительно увеличить точность распознавания.

***Благодарности.** Работа выполнена в рамках программы «Приоритет-2030», выполняемой КНИТУ-КАИ (молодые ученые).*

### Литература

1. Wu W., Liu Z., Chen M. An automated vision system for container-code recognition // Expert Systems with Applications. 2012. V. 39. URL: doi.org/10.1016/j.eswa.2011.08.143.
2. Yoon Y., Yoon H., Kim D. Automatic Container Code Recognition From Multiple Views // Etri Journal. 2016. V. 38. URL: doi.org/10.4218/etrij.16.0014.0069.
3. Chun-ming T., Peng C. Container Number Recognition Method Based on SSD\_MobileNet and SVM // American Academic Scientific Research Journal for Engineering, Technology, and Sciences. 2020. V. 74. pp. 200-211.
4. Zhou X., Yau C., Wen H. EAST: An Efficient and Accurate Scene Text Detector // Computer Vision and Pattern Recognition, 2017. URL: arxiv.org/abs/1704.03155v2.
5. Макаров Р.А. Распознавание цифр маркировки грузового контейнера с использованием алгоритма Faster-RCNN // Инженерный вестник Дона, 2023, №3. URL: ivdon.ru/ru/magazine/archive/n3y2023/8276.
6. Евсина В.А., Широбокова С.Н., Жжонов В.А., Евсин В.А. Моделирование и реализация процесса распознавания знаков дорожного

движения при определении ситуации на дороге с помощью искусственных нейронных сетей // Инженерный вестник Дона, 2022, №4. URL: [ivdon.ru/ru/magazine/archive/n4y2022/7587](http://ivdon.ru/ru/magazine/archive/n4y2022/7587)

7. Николенко С.И., Кадурин А.А., Архангельская Е.В. Глубокое обучение. Погружение в мир нейронных сетей. СПб.: Питер, 2018. 480 с.

8. Nigar N., Faisal H.M., Shahzad M.K. An Offline Image Auditing System for Legacy Meter Reading Systems in Developing Countries: A Machine Learning Approach // Journal of Electrical and Computer Engineering. 2022. V. 2022. URL: [doi.org/10.1155/2022/4543530](https://doi.org/10.1155/2022/4543530).

9. Бережнов Н.И., Сирота А.А. Универсальный алгоритм улучшения изображений с использованием глубоких нейронных сетей. // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2022. №2. С. 81-92.

10. Ren, Y.; Zhu, C.; Xiao, S. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN // Appl. Sci., 2018, №8. DOI: [doi.org/10.3390/app8050813](https://doi.org/10.3390/app8050813).

11. Гизатуллин З.М., Ляшева С.А., Морозов О.Г., Шлеймович М.П. Метод обнаружения контуров на основе весовой модели изображения. // Компьютерная оптика. 2022. Т.44. № 3. С. 393-400.

12. Ляшева С.А., Морозов О.Г., Шлеймович М.П. Анализ весовой модели для сжатия изображений на основе вейвлет-преобразования. // Прикаспийский журнал: управление и высокие технологии. 2020. № 3 (51). С. 9-22.

13. Lyasheva S., Morozov O., Shleymovich M. Analysis of energy characteristics for issuing areas of significance when compressing images in cyber-physical systems. // Studies in Systems, Decision and Control. 2021. Т. 350. pp. 259-270.

## References

1. Wu W., Liu Z., Chen M. Expert Systems with Applications. 2012. V. 39. URL: [doi.org/10.1016/j.eswa.2011.08.143](https://doi.org/10.1016/j.eswa.2011.08.143).
2. Yoon Y., Yoon H., Kim D. Etri Journal. 2016. V. 38. URL: [doi.org/10.4218/etrij.16.0014.0069](https://doi.org/10.4218/etrij.16.0014.0069).
3. Chun-ming T., Peng C. American Academic Scientific Research Journal for Engineering, Technology, and Sciences. 2020. V. 74. pp. 200-211.
4. Zhou X., Yau C., Wen H. Computer Vision and Pattern Recognition, 2017. URL: [arxiv.org/abs/1704.03155v2](https://arxiv.org/abs/1704.03155v2).
5. R.A. Makarov Inzhenernyj vestnik Dona, 2023, №3. URL: [ivdon.ru/ru/magazine/archive/n3y2023/8276](https://ivdon.ru/ru/magazine/archive/n3y2023/8276).
6. V.A. Evsina, S.N. Shirobokova, V.A. Zhzhonov, V.A. Evsin Inzhenernyj vestnik Dona, 2022, №4. URL: [ivdon.ru/ru/magazine/archive/n4y2022/7587](https://ivdon.ru/ru/magazine/archive/n4y2022/7587).
7. Nikolenko S.I., Kadurin A.A., Arhangel'skaya E.V. Glubokoe obuchenie. Pogruzhenie v mir nejronnyh setej [Deep learning. Loading into the world of neural networks]. Saint Petersburg: Piter, 2018. 480 p.
8. Nigar N., Faisal H.M., Shahzad M.K. Journal of Electrical and Computer Engineering. 2022. V. 2022. URL: [doi.org/10.1155/2022/4543530](https://doi.org/10.1155/2022/4543530).
9. Berezhnov N.I., Sirota A.A. Vestnik Voronezskogo gosudarstvennogo universiteta. Seria Sistemnyj analiz i informacionnye tehnologii. 2022. №2. pp. 81-92.
10. Ren, Y.; Zhu, C.; Xiao, S. Appl. Sci., 2018, №8. DOI: [doi.org/10.3390/app8050813](https://doi.org/10.3390/app8050813)
11. Gizatullin Z.M., Lyasheva S.A., Morozov O.G., Shleymovich M.P. Komp'uternaa optika. 2022. T.44. № 3. pp. 393-400.
12. Lyasheva S.A., Morozov O.G., Shleymovich M.P. Prikaspijskij zurnal: upravlenie i vysokie tehnologii. 2020. № 3 (51). pp. 9-22.



13. Lyasheva S., Morozov O., Shleymovich M. Studies in Systems, Decision and Control. 2021. T. 350. pp. 259-270.