**УДК 004.896**

**Применение глубокого обучения для аугментации и генерации подводного набора данных с промышленными объектами**

А. Ахмад1,2,3,Н.А. Андриянов3, В.И. Соловьев2,4, Д.А. Соломатин3

1ФГБОУ ВО «Московский государственный технический университет им. Н.Э. Баумана», г. Москва, Россия

2ООО «ЦИАРС», г. Москва, Россия

3ФГБОУ ВО «Финансовый университет при Правительстве Российской Федерации», г. Москва, Россия

4ФГБОУ ВО «Московский технический университет связи и информатики», г. Москва, Россия

**Аннотация**

*Цель исследования:* разработка метода глубокого обучения для аугментации и генерации проблемно-ориентированного набора данных, содержащего промышленные объекты.

*Методы исследования:* архитектура CycleGAN, обеспечивающая преобразование набора данных, содержащего изображения различных объектов, сделанные в лабораторной или в обычной надземной среде, в набор данных, содержащий характеристики подводной среды.

*Результаты исследования:* представлен инструмент для решения проблемы отсутствия подводных наборов данных, разработана модель глубокого обучения, которая применяется для создания изображений с подводными элементами.

*Заключение:* исследование показало, что модель CycleGAN можно использовать для создания изображений различных объектов в подводной среде.

***Ключевые слова:*** автономные необитаемые подводные аппараты (АНПА), машинное обучение; глубокое обучение; CNN (сверточные нейронные сети); состязательные потери (adversarial losses); CycleGAN; дискриминатор

**Application of deep learning for augmentation and generation of an underwater data set with industrial facilities**

A. Ahmad1,2,3, N.A. Andriyanov2, V.I. Soloviev, D.A. Solomatin2

1Bauman Moscow State Technical University, Moscow, Russia

2CIARS Llc, Moscow, Russia

3Financial University under the Government of the Russian Federation, Moscow, Russia

4Moscow Technical University of Communications and Informatics, Moscow, Russia

**Abstract**

*The purpose of the study*: development of a deep learning method for augmentation and generation of a problem-oriented data set containing industrial objects.

*Research methods*: CycleGAN architecture that converts a dataset containing images of various objects taken in a laboratory or in a conventional above-ground environment into a dataset containing characteristics of the underwater environment.

*Results of the study*: a tool is presented to solve the problem of the lack of underwater data sets, a deep learning model is developed, which is used to create images with underwater elements.

*Conclusion*: the study showed that the CycleGAN model can be used to create images of various objects in the underwater environment.

***Keywords***: autonomous underwater vehicles (AUV), machine learning; deep learning; CNN (Convolutional neural networks); adversarial losses; CycleGAN; discriminator

**Введение**

В настоящее время растет потребность и интерес к использованию и разработке автономных необитаемых подводных аппаратов (АНПА) для изучения и защиты подводной среды, проведения поисково-спасательных работ, разведывательных операций и операций по обнаружению объектов в морской среде. Особенно такая потребность важна в тех средах, в которых содержатся промышленные объекты: нефте- и газопроводы, ветряные турбины, интернет-кабели и т.д.

Из-за сложности доступа к этим участкам, дороговизны и опасности для человека при проведении этих операций, научное сообщество приступило к разработке автономных подводных аппаратов, способных выполнять такие задачи автономно, без участия человека [1-3].

Обнаружение объектов является важной частью рабочего процесса обработки изображений. Хотя существует множество подходов, основанных на компьютерном зрении [4-6], их производительность остается достаточно низкой из-за особенностей подводной среды. Во многом это связано с препятствиями, которые рассеивают и поглощают свет в подводных условиях.

После научной революции в области глубокого обучения большое внимание стало уделяться внедрению этих технологий в подводные операции [7-9], чтобы робот мог выполнять эти задачи автоматически. Технологии глубокого обучения позволили значительно продвинуться в цифровой обработке изображений для распознавания и категоризации объектов. Использование глубокого обучения позволило решить широкий круг вопросов, включая защиту морской экосистемы, спасение жизней в чрезвычайных ситуациях, предотвращение подводных бедствий, а также обнаружение, отслеживание и идентификацию подводных целей.

Но задачи, связанные с обнаружением подводным объектов, сталкиваются с трудностями из-за нехватки обучающих наборов данных. Исследователи решают эту проблему, используя новые технологии, ставшие доступными благодаря быстрому развитию искусственного интеллекта в последние годы.

# Задача обнаружения объектов в подводной среде

Подводные автономные роботы должны понимать сложную окружающую среду в большинстве миссий, чтобы проводить полностью автономные операции. При этом обнаружение объектов является одной из основных задач разведки и поисково-спасательных операций. Сложная подводная среда создает трудности для обнаружения объектов: несбалансированные условия освещения, низкая контрастность, перекрытие объектами друг друга. Изображения, получаемые с подводной камеры, как правило, являются расплывчатыми, и обычные детекторы компьютерного зрения часто не справляются с обнаружением объектов на таких изображения. Появившиеся в последние несколько лет модели машинного обучения, став более сложными и точными, позволяют обнаруживать и сегментировать объекты в том числе и в таких сложных условиях. Некоторые знаковые нейросетевые архитектуры, которые произвели эту революцию, включают семейство моделей R-CNN [10], YOLO [11], FPN [12] и др.

# Наборы данных в подводной среде

Машинное обучение (Machine Learning, ML) — это набор методов анализа данных, который позволяет обучать аналитические системы решать повторяющиеся типовые аналитические задачи без использования программирования, на основе выявления закономерностей или скрытых паттернов в данных и принятия решений с минимальным участием человека.

Для обучения требуется размеченный набор данных – обучающие данные, которые содержат примеры решения проблем. Алгоритмы машинного обучения выявляют в них закономерности и в результате могут прогнозировать результат на неразмеченных данных с контролируемым качеством прогноза, поэтому их можно использовать на практике.

При применении подходов глубокого обучения в нашей задаче, связанной с классификацией или обнаружением водолазов и посторонних объектов в подводной среде, одним из наиболее частых препятствий является нехватка обучающих наборов данных.

На сегодняшний день доступно лишь небольшое количество наборов изображений подводной среды для обучения моделей компьютерного зрения [13]. Известные наборы изображений перечислены в Таблице 1.

**Таблица 1. Наборы изображений подводной среды**

**Table 1. Underwater environment image sets**

|  |  |
| --- | --- |
| MUED [14] | Marine Underwater Environment Database – набор 8600 подводных изображений объектов 430 классов со сложным фоном и различными вариациями пространственного положения, освещения, мутности воды и т.д. |
| RUIE [15] | Real-time Underwater Image – набор данных, состоящий из трех поднаборов: Underwater Image Quality Set (UIQS), Underwater Color Cast Set (UCCS), Underwater Task-driven Testset (UTTS) |
| SUIM [16] | Segmentation of Underwater Imagery – набор данных, содержащий более 1500 изображений объектов 8 категорий: рыб, рифов, водных растений, затонувших кораблей, дайверы, роботы и морское дно |
| TrashCan [17] | Набор данных для сегментации экземпляров подводного мусора, содержащий 7212 изображений подводных роверов, подводного мусора, а также подводной флоры и фауны |
| UIEB [18] | Underwater Image Enhancement Benchmark Dataset – 950 реальных подводных изображений, 890 из которых имеют соответствующие эталонные изображения |
| Underwater\_ImageNet [18, 19] | Набор подводных объектов, вручную отобранных из набора ImageNet, а также из подводных видео с Youtube |
| Test-lhnog [20] | Набор изображений экипировки аквалангистов |
| Human [21] | Набор изображений аквалангистов |
| Please [22] | Набор изображений потерпевших крушение кораблей |
| P [23] | Набор изображений производственных труб |

Проблемы нехватки данных и малое количество наборов изображений с подводной средой обусловлены следующими причинами:

* сбору подводных изображений не уделялось достаточного внимания;
* возникали трудности, связанные с подводной средой, поскольку сбор подводных данных – трудоемкая работа, требующая привлечения специалистов и использования специального оборудования.

Для решения проблемы нехватки данных обычно применяют методы аугментации данных (data augmentation) [24, 25]. Классические методы аугментации включают внесение изменений в набор данных путем применения таких операций, как вращение и масштабирование. Эти методы часто используются для увеличения размера набора данных и его разнообразия. Неклассические методы аугментации включают внесение изменений в набор данных путем применения более продвинутых методов, таких как добавление шума, обрезка и изменение цвета изображений [26].

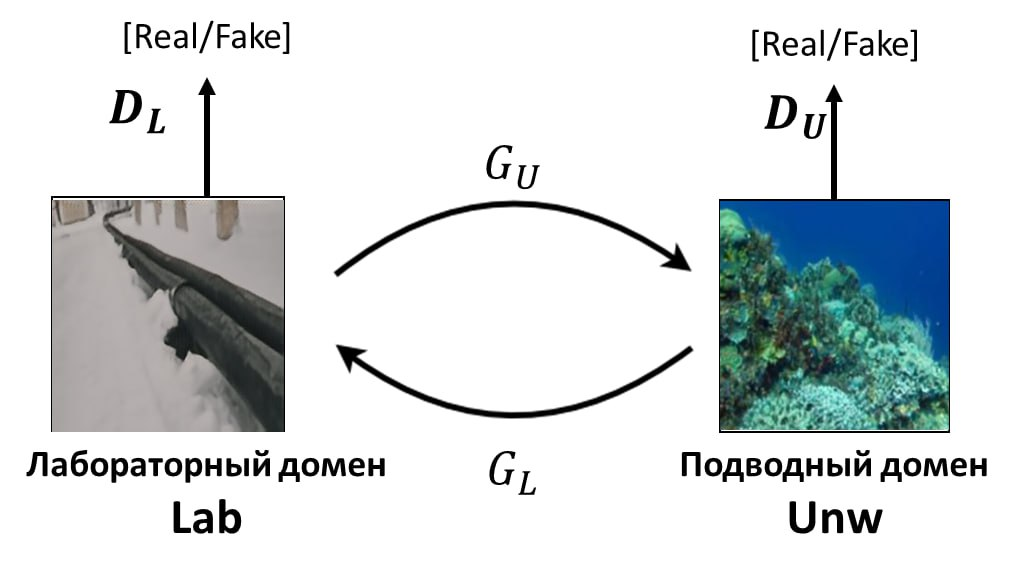
Для таких задач, как перенос стиля, т. е. фоновой составляющей желаемой среды, классические методы аугментации данных не могут генерировать изображения, близкие к реальным. К тому же, как было отмечено, наборов изображений промышленных объектов (например, нефте- и газопроводов) существует крайне мало, и их тяжело собрать.

Для решения нам потребуются такие модели глубокого обучения, как генеративно-состязательные сети (Generative Adversarial Networks GAN) [27], CycleGAN [28] и U-Net [29], которые являются современными методами, используемыми для расширения наборов данных и увеличения их размера [30, 31]. Генеративно-состязательные сети в основном используются для создания синтетических изображений, которые следуют тому же распределению вероятностей, что и реальные изображения. CycleGAN — это хорошо известная архитектура генеративно-состязательных сетей, которая обычно используется для изучения преобразований изображений по различным шаблонам.

# CycleGAN

Cycle Generative Adversarial Network (CycleGAN) — это подход к обучению глубоких сверточных сетей (Convolutional Neural Networks CNN) [10] для задач преобразования одного изображения в другое. В отличие от других моделей GAN для задач перевода изображений, CycleGAN изучает отображения (mappings) между одним доменом изображения и другим, используя обучение без учителя (unsupervised learning).

Нас интересует преобразование изображения целевых объектов, которые находятся в лабораторном домене (Lab), в изображение в домене подводной среды (Unw). Способ, которым CycleGAN [28] осуществляет такое преобразование, заключается в обучении сетей генераторов отображению из домена Lab в изображение, которое выглядит так, как будто оно получено из домена Unw (и наоборот), как показано на Рис. 1.



**Рис. 1. Схема работы CycleGAN**

**Fig. 1. CycleGAN scheme**

Отобразим задачи генераторов, как перевод среди множеств подводной и лабораторной среды

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

и — это генераторы, которые берут изображение из одного домена и переводят его в другой. отображает картинку из в Lab (), тогда как идет в противоположном направлении, отображая из в  ().

У каждого генератора есть соответствующий дискриминатор, который пытается отличить синтезированные изображения от реальных.

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

дискриминатор обеспечивает состязательное обучение (adversarial learming) для , а делает то же самое для .

Оба генератора и обучаются «обману» соответствующего дискриминатора, чтобы он был менее способен отличать сгенерированные изображения от реальных версий в каждом домене.

Модели дискриминатора и генератора обучаются в состязательном процессе, как и обычные модели GAN [27]. Данный процесс обучения предполагает оптимизацию двух типов потерь:

1. Adversarial Losses – состязательные потери для сопоставления распределения сгенерированных изображений с распределением данных в целевом домене (3):

|  |  |
| --- | --- |
|  | (3) |

где ;

1. Cycle Consistency Losses – потери согласованности цикла для предотвращения противоречия изученных отображений и друг другу (4):

|  |  |
| --- | --- |
|  | (4) |

Потери согласованности цикла иллюстрирует Рис. 2.

Теперь можно записать полную целевую функцию, объединив эти условия потерь в свертку и присвоив потере согласованности цикла вес λ:

|  |  |
| --- | --- |
|  | (5) |

Значение гиперпараметра подбирается как решение задачи минимизации функции потерь (5).

Graphical user interface

Description automatically generated

**Рис. 2. Потери согласованности цикла**

**Fig. 2. Cycle consistency losses**

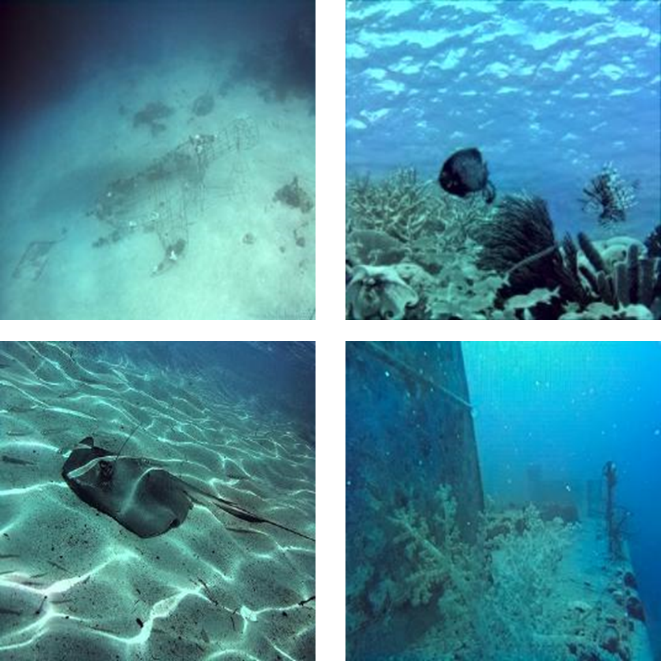
## Наборы данных CycleGAN

Из-за перечисленных особенностей подводной среды сложно получить набор данных, содержащий изображения деталей механического и промышленного оборудования, которые потенциально могут быть обнаружены в подводных структурах.

Для генерации набора данных с целевыми объектами в подводной среде в данной работе используется подход перевода изображений в изображения CycleGAN (CycleGAN Image-to-Image Translation) [28, 31].

Вначале из общедоступных наборов данных был отобран набор изображений подводной среды, представляющих ее наиболее точно (набор данных Underwater\_ImageNet) [19] (Рис. 3). Затем был собран базовый набор данных – набор изображений объектов механического и промышленного оборудования, которые потенциально могут быть обнаружены в подводной среде [32] (Рис. 4). Данный этап выполнен в лабораторных условиях.

На следующем этапе набор изображений подводной среды и базовый набор изображений были использованы для обучения модели CycleGAN.



**Рис. 3 Набор изображений подводной среды (underwater\_ImageNet) [19]**

**Fig. 3. Underwater environment image set (underwater\_ImageNet) [19]**



**Рис. 4. Базовый набор изображений [32]**

**Fig. 4. The basic image set [32]**

## 

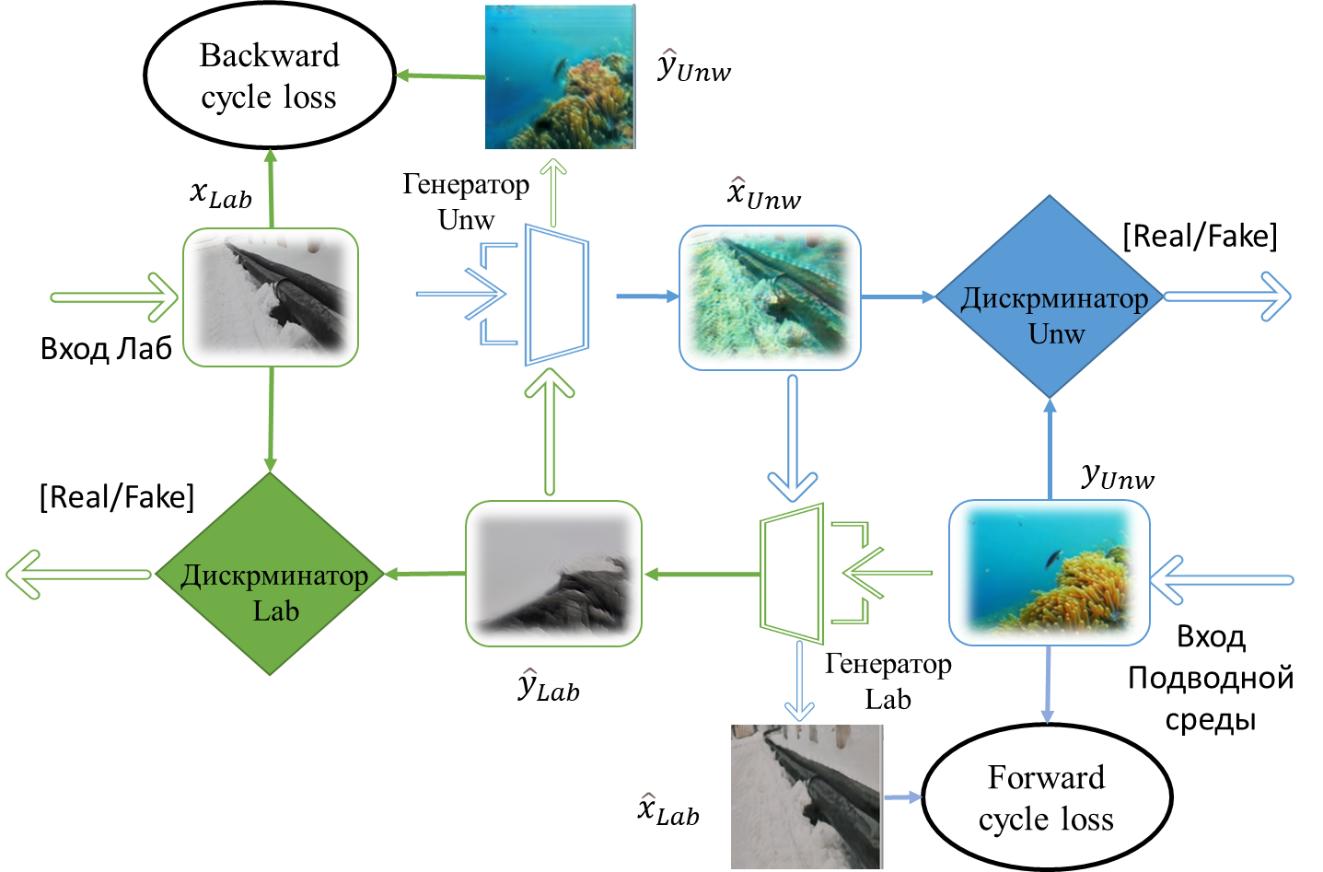
## Архитектура модели CycleGAN

Архитектура модели CycleGAN представлена на Рис 5. В этой модели характеристики изображения подводного домена преобразуются в лабораторный домен с помощью генератора .

Затем генератор преобразует изображения лабораторного домена в подводный домен, что в итоге приводит к созданию искусственных изображений .

Дискриминаторы отвечают за мониторинг и сравнение искусственных изображений в каждом домене с подлинными на самой последней фазе сети, при этом вычисляются состязательные подери (3), по которым оптимизируется окончательный вывод изображений дискриминаторами.

Затем на каждом этапе процесса модель реконструирует искусственные изображения в исходный домен и рассчитывает потери согласованности цикла (4).



**Рис. 5. Общая схема CycleGAN**

**Fig. 5. The general scheme of CycleGAN**

# 

# Результаты

Основная цель глубокого обучения для аугментации данных состоит в генерации изображений механического и промышленного оборудования в подводной среде.

На рис. 6 показаны некоторые результаты, полученные с помощью модели CycleGAN. Видно, что модель генерирует новые изображения так, как будто они действительно были сделаны в подводной среде.

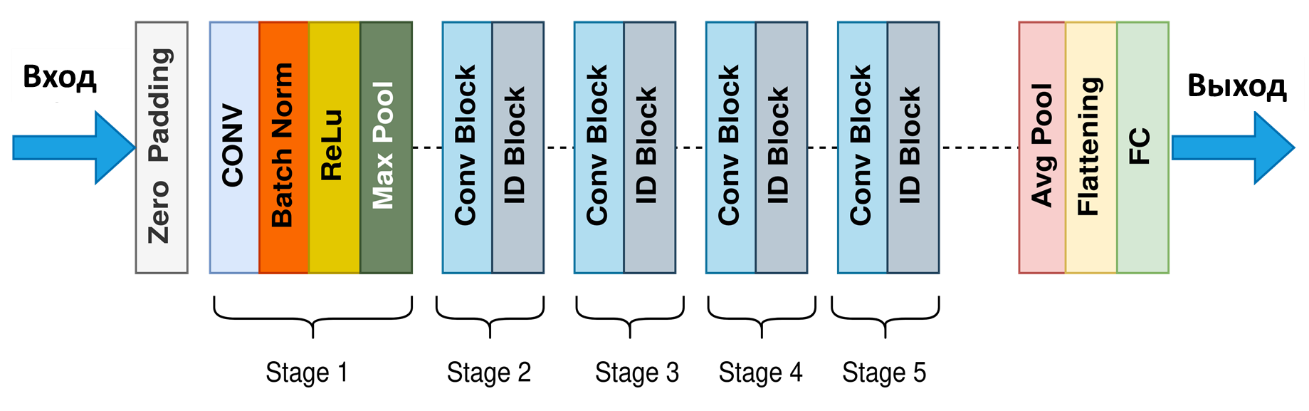


**Рис. 6. Примеры сгенерированных изображений**

**Fig. 6. Generated images examples**

Для проверки качества результатов генерации подводных изображений был использован классификатор сгенерированных изображений – предобученная модель ResNet-50 [33]. ResNet-50 (Residual Network – остаточная сеть) представляет собой 50-слойную сверточную нейронную сеть (48 сверточных слоев, один слой MaxPool и один средний слой пула), в которой проблема исчезающих градиентов решается с использованием концепции быстрых подключений, которые «пропускают» некоторые уровни, преобразуя обычную сеть в остаточную (Residual).

На рис. 7 представлена архитектура ResNet-50, где был отредактирован последний полносвязный слой **,** и установлен выход на 2 класса. Был обучен только последний линейный слой с использованием 3000 изображений подводной среды (из набора Underwater\_ImageNet) и 500 изображений из базового набора.

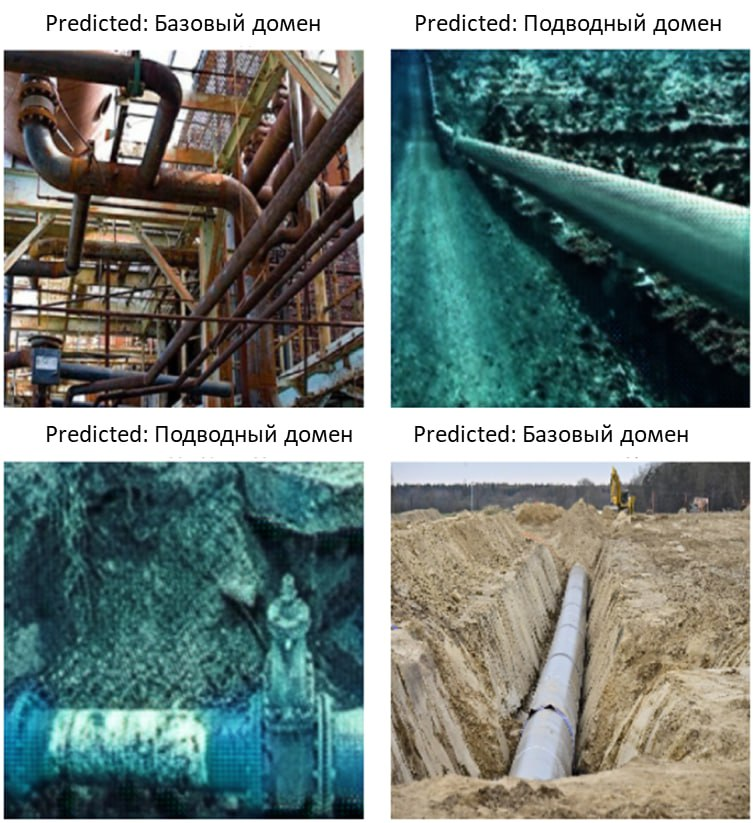


**Рис. 7. Архитектура ResNet-50**

**Fig. 7. ResNet-50 architecture**

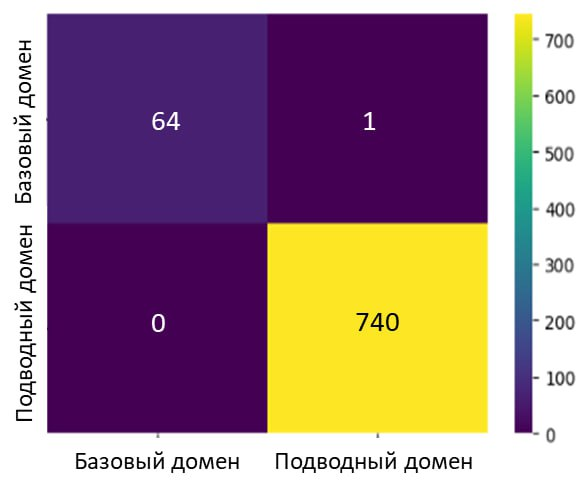
Примеры результатов классификации сгенерированных изображений представлены на Рис. 8. При этом все изображения, созданные с помощью CycleGAN, как изобыли классифицированы как изображения в подводном домене.

Матрица ошибок классификации, представленная на Рис. 9, позволяет оценить точность классификации полноту и метрику



**Рис. 8. Примеры классификации изображений**

**Fig. 8. Image classification examples**



**Рис. 9. Матрица ошибок классификации**

**Fig. 8. Classification confusion matrix**

# Заключение

В данной работе представлен инструмент для решения проблемы отсутствия наборов изображений промышленного оборудования в подводной среде, актуальной в связи с ростом потребности в ремонтных, исследовательских, спасательных и других подводных работах в подводной среде и отсутствием инструмента для создания искусственных подводных изображений.

Была разработана модель глубокого обучения, которая применяется для создания изображений с подводными элементами. В модели используются сделанные в надводной среде изображения объектов, которые потенциально можно обнаружить под водой (труб, клапанов, фланцев, винтов и т.п.) и изображения подводной среды, взятые из общедоступных наборов данных. Исследование показало, что сверточная нейронная сеть CycleGAN способна эффективно генерировать изображения объектов в подводной среде ( на тестовой выборке из изображений).

Сгенерированный новый набор изображений можно использовать для решения задач сегментации изображений и обнаружения объектов с помощью трансферного обучения [34–37].

**Литература**

1. Ridolfi, A., Conti, R., Costanzi, R., et al. A dynamic manipulation strategy for an intervention: Autonomous underwater vehicle // Advances in Robotics and Automation. – 2015. – V. 4. – No. 2. – Article 1000132. – DOI: 10.4172/2168-9695.1000132
2. Wynn R.B., Huvenne V.A.I., Le Bas T.P., et al. Autonomous underwater vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience // Marine Geology. – 2014. – V. 352. – P. 451–468. – DOI: 10.1016/j.margeo.2014.03.012
3. Alam K., Ray T., Anavatti S.G. A brief taxonomy of autonomous underwater vehicle design literature // Ocean Engineering. – 2014. – V. 88. –  
   P. 627–630. – DOI: 10.1016/j.oceaneng.2014.04.027
4. Manzanilla A., Reyes S., Garcia M., et al. Autonomous navigation for unmanned underwater vehicles: Real-time experiments using computer vision // IEEE Robotics and Automation Letters. – 2019. – V. 4. – No. 2. – P. 1351–1356. – DOI: 10.1109/LRA.2019.2895272
5. Reggiannini M., Moroni D. The use of saliency in underwater computer vision: A review // Remote Sensing. – 2021. – V. 13. – No. 1. – Article 22. – DOI: 10.3390/rs13010022
6. Andriyanov N.A., Dementiev V.E., Tashlinskii A.G. Detection of objects in the images: From likelihood relationships towards scalable and efficient neural networks // Computer Optics. – 2022. – V. 46. – No. 1. – P. 139–159. – DOI: 10.18287/2412-6179-CO-922.
7. Jin L., Liang H. Deep learning for underwater image recognition in small sample size situations // OCEANS 2017-Aberdeen. – 2017. – P. 1–4. – DOI: 10.1109/OCEANSE.2017.8084645
8. Meng L., Hirayama T., Oyanagi S. Underwater-drone with panoramic camera for automatic fish recognition based on deep learning // IEEE Access. – 2018. – V. 6. – P. 17880–17886. – DOI: 10.1109/ACCESS.2018.2820326
9. Yang H., Liu p., Hu Y., Fu J. Research on underwater object recognition based on YOLOv3 // Microsystem Technologies. – 2021. – V. 27. – No. 9. – P. 1837–1844. – DOI: 10.1007/s00542-019-04694-8
10. Girshick R. Fast R-CNN // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2015. – P. 1440-1448. – arXiv:1504.08083
11. Bochkovskiy A., Wang C.Y., Liao H.Y.M. YOLOv4: Optimal speed and accuracy of object detection. – 2020. – arXiv:2004.10934.
12. Cazzato D. et al. A survey of computer vision methods for 2d object detection from unmanned aerial vehicles // Journal of Imaging. – 2020. – V. 6. – No. 8. – Article 78. – DOI: 10.3390/jimaging6080078
13. Dakhil R.A., Khayeat A.R.H. Review on deep learning technique for underwater object detection // Proceedings of the 3rd International Conference on Data Science and Machine Learning (DSML). – 2022. – P. 49–63. – arXiv:2209.10151
14. Jian M., Qi Q., Yu H., et al. The extended marine underwater environment database and baseline evaluations // Applied Soft Computing. – 2019. – V. 80. – P. 425–437. – DOI: 10.1016/j.asoc.2019.04.025
15. Liu R., Hou M., Fan X., et al. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light // IEEE Transactions on Circuits and Systems for Video Technology. – 2020. – V. 30. – No.12. – P. 4861–4875. – DOI: 10.1109/TCSVT.2019.2963772
16. Islam M.J., Edge C., Xiao Y., et al. Semantic segmentation of underwater imagery: Dataset and benchmark // Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). – 2020. – DOI: 10.1109/IROS45743.2020.9340821
17. Hong J., Fulton M., Sattar J. TrashCan: A semantically-segmented dataset towards visual detection of marine debris. – 2020. – arXiv:2007.08097
18. Fabbri C., Islam M.J., Sattar J. Enhancing underwater imagery using generative adversarial networks // Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). – 2018. – P. 7159–7165. – arXiv:1801.04011
19. Набор данных Underwater\_ImageNet. – 2022. –<https://drive.google.com/file/d/1LOM-2A1BSLaFjCY2EEK3DA2Lo37rNw-7/view> (дата посещения: 07.03.22).
20. Набор данных Test-lhnog. // Roboflow Universe. – 2022. –<https://universe.roboflow.com/new-workspace-qnaxk/test-lhnog> (дата посещения: 07.03.23).
21. Набор данных Human // Roboflow Universe. – 2022. – <https://universe.roboflow.com/school-9j0uz/human-ighgb> (дата посещения: 07.03.23).
22. Набор данных Please // Roboflow Universe. – 2022. – <https://universe.roboflow.com/new-workspace-bpkjt/please> (дата посещения: 07.03.23).
23. Набор данных P // Roboflow Universe. – 2023. – <https://universe.roboflow.com/pipes-nty6m/p-w1typ> (дата посещения: 07.03.23).
24. Perez L., Wang J. The effectiveness of data augmentation in image classification using deep learning. – 2017. – arXiv:1712.04621
25. Andriyanov N.A., Andriyanov D.A. The using of data augmentation in machine learning in image processing tasks in the face of datascarity // Journal of Physics Conference Series. – 2020. – V. 1661. – No. 1. – Article 012018. – DOI: 10.1088/1742-6596/1661/1/012018
26. Buslaev A., Iglovikov V.I., Khvedchenya E., et al. Albumentations: Fast and flexible image augmentations // Information. – 2020. – V. 11. – No. 2. – Article 125. – DOI: 10.3390/info11020125
27. Goodfellow I.J., Pouget-Abadie J., Mirza M., et al. Generative adversarial networks // Communications of the ACM. – 2020. – V. 63. – No. 1. – P. 139–144. – DOI: 10.1145/3422622
28. Zhu J.Y., Park T., Isola P., Efros A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 2223–2232. – arXiv:1703.10593
29. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. – 2015. – arXiv:1505.04597
30. Polymenis I., Haroutunian M., Norman R., Trodden D. Virtual underwater datasets for autonomous inspections // Journal of Marine Science and Engineering. – 2022. – V. 10. – No. 9. – Article 1289. – DOI: 10.3390/jmse10091289
31. Welander P., Karlsson S., Eklund A. Generative adversarial networks for image-to-image translation on multi-contrast MR images – A comparison of CycleGAN and UNIT. – 2018. – arXiv:1806.07777
32. Набор изображений механического и промышленного оборудования / Сост. А. Ахмад. – 2022. – <https://drive.google.com/drive/folders/1p4mSIse4PEPx9pDR7TMDgL6nPVnLpzFi?usp=share_link> (дата посещения: 07.03.23).
33. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 770–778. – arXiv:1512.03385
34. Andriyanov N.A., Dementiev V.E., Tashlinskiy A.G. Deep Markov models of multidimensional random fields // Procedia Computer Science. – 2020. – V. 176. – P. 1289–1298. – DOI: 10.1016/j.procs.2020.09.138
35. Andriyanov N., Khasanshin I., Utkin D., et al. Intelligent system for estimation of the spatial position of apples based on YOLOv3 and Real Sense depth camera D415 // Symmetry. – 2022. – V. 14. – No. 1. – Article 148. – DOI: 10.3390/sym14010148
36. Кузнецова А.А., Малева Т.В., Соловьев В.И. Использование алгоритма YOLOv3 с процедурами пред- и постобработки для обнаружения плодов роботом для сбора яблок // Прикладная информатика. – 2020. – Т. 15. – № 4. – С. 64–74. – DOI: 10.37791/2687-0649-2020-15-4-64-74
37. Калашников В.А., Соловьев В.И. Приложения компьютерного зрения в горнодобывающей промышленности // Прикладная информатика. 2023. – Т. 18. – № 1. – С. 4–21. – DOI: 10.37791/2687-0649-2023-18-1-4-21

**References**

1. Ridolfi, A., Conti, R., Costanzi, R., et al. A dynamic manipulation strategy for an intervention: Autonomous underwater vehicle, Advances in Robotics and Automation, 2015, Vol. 4, No. 2, article 1000132, DOI: 10.4172/2168-9695.1000132
2. Wynn R.B., Huvenne V.A.I., Le Bas T.P., et al. Autonomous underwater vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience, Marine Geology, 2014, Vol. 352, pp. 451-468, DOI: 10.1016/j.margeo.2014.03.012
3. Alam K., Ray T., Anavatti S.G. A brief taxonomy of autonomous underwater vehicle design literature, Ocean Engineering, 2014, Vol. 88.  
   P. 627-630, DOI: 10.1016/j.oceaneng.2014.04.027
4. Manzanilla A., Reyes S., Garcia M., et al. Autonomous navigation for unmanned underwater vehicles: Real-time experiments using computer vision, IEEE Robotics and Automation Letters, 2019, Vol. 4, No. 2, pp. 1351-1356, DOI: 10.1109/LRA.2019.2895272
5. Reggiannini M., Moroni D. The use of saliency in underwater computer vision: A review, Remote Sensing, 2021, Vol. 13, No. 1, article 22, DOI: 10.3390/rs13010022
6. Andriyanov N.A., Dementiev V.E., Tashlinskii A.G. Detection of objects in the images: From likelihood relationships towards scalable and efficient neural networks, Computer Optics, 2022, Vol. 46, No. 1, pp. 139-159, DOI: 10.18287/2412-6179-CO-922.
7. Jin L., Liang H. Deep learning for underwater image recognition in small sample size situations, OCEANS 2017-Aberdeen, 2017, pp. 1-4, DOI: 10.1109/OCEANSE.2017.8084645
8. Meng L., Hirayama T., Oyanagi S. Underwater-drone with panoramic camera for automatic fish recognition based on deep learning, IEEE Access, 2018, Vol. 6, pp. 17880-17886, DOI: 10.1109/ACCESS.2018.2820326
9. Yang H., Liu p., Hu Y., Fu J. Research on underwater object recognition based on YOLOv3, Microsystem Technologies, 2021, Vol. 27, No. 9, pp. 1837-1844, DOI: 10.1007/s00542-019-04694-8
10. Girshick R. Fast R-CNN, Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, P. 1440-1448. – arXiv:1504.08083
11. Bochkovskiy A., Wang C.Y., Liao H.Y.M. YOLOv4: Optimal speed and accuracy of object detection, 2020, arXiv:2004.10934.
12. Cazzato D. et al. A survey of computer vision methods for 2d object detection from unmanned aerial vehicles, Journal of Imaging, 2020, Vol. 6, No. 8, article 78, DOI: 10.3390/jimaging6080078
13. Dakhil R.A., Khayeat A.R.H. Review on deep learning technique for underwater object detection, Proceedings of the 3rd International Conference on Data Science and Machine Learning (DSML), 2022, pp. 49-63, arXiv:2209.10151
14. Jian M., Qi Q., Yu H., et al. The extended marine underwater environment database and baseline evaluations, Applied Soft Computing, 2019, Vol. 80, pp. 425-437. - DOI: 10.1016/j.asoc.2019.04.025
15. Liu R., Hou M., Fan X., et al. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light, IEEE Transactions on Circuits and Systems for Video Technology, 2020, Vol. 30, No.12, pp. 4861-4875, DOI: 10.1109/TCSVT.2019.2963772
16. Islam M.J., Edge C., Xiao Y., et al. Semantic segmentation of underwater imagery: Dataset and benchmark, Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020, DOI: 10.1109/IROS45743.2020.9340821
17. Hong J., Fulton M., Sattar J. TrashCan: A semantically-segmented dataset towards visual detection of marine debris, 2020, arXiv:2007.08097
18. Fabbri C., Islam M.J., Sattar J. Enhancing underwater imagery using generative adversarial networks, Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 7159-7165, arXiv:1801.04011
19. Underwater\_ImageNet Dataset, 2022, <https://drive.google.com/file/d/1LOM-2A1BSLaFjCY2EEK3DA2Lo37rNw-7/view> (Accessed 07 March 23).
20. Test-lhnog Dataset, Roboflow Universe, 2022, <https://universe.roboflow.com/new-workspace-qnaxk/test-lhnog> (Accessed 07 March 23).
21. Human Dataset, Roboflow Universe, 2022, <https://universe.roboflow.com/school-9j0uz/human-ighgb> (Accessed 07 March 23).
22. Please Dataset, Roboflow Universe, 2022, <https://universe.roboflow.com/new-workspace-bpkjt/please> (Accessed 07 March 23).
23. P Dataset, Roboflow Universe, 2023, <https://universe.roboflow.com/pipes-nty6m/p-w1typ> (Accessed 07 March 23).
24. Perez L., Wang J. The effectiveness of data augmentation in image classification using deep learning, 2017, arXiv:1712.04621
25. Andriyanov N.A., Andriyanov D.A. The using of data augmentation in machine learning in image processing tasks in the face of datascarity, Journal of Physics Conference Series, 2020, Vol. 1661, No. 1, article 012018, DOI: 10.1088/1742-6596/1661/1/012018
26. Buslaev A., Iglovikov V.I., Khvedchenya E., et al. Albumentations: Fast and flexible image augmentations, Information, 2020, Vol. 11, No. 2, article 125, DOI: 10.3390/info11020125
27. Goodfellow I.J., Pouget-Abadie J., Mirza M., et al. Generative adversarial networks, Communications of the ACM, 2020, Vol. 63, No. 1, pp. 139-144, DOI: 10.1145/3422622
28. Zhu J.Y., Park T., Isola P., Efros A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks, Proceedings of the IEEE International Conference on Computer Vision (ICCV),2017, pp. 2223-2232, arXiv:1703.10593
29. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation, 2015, arXiv:1505.04597
30. Polymenis I., Haroutunian M., Norman R., Trodden D. Virtual underwater datasets for autonomous inspections, Journal of Marine Science and Engineering, 2022, Vol. 10, No. 9, article 1289, DOI: 10.3390/jmse10091289
31. Welander P., Karlsson S., Eklund A. Generative adversarial networks for image-to-image translation on multi-contrast MR images – A comparison of CycleGAN and UNIT, 2018, arXiv:1806.07777
32. Mechanical and industrial equipment image set / Comp. by A. Ahmad, 2022, <https://drive.google.com/drive/folders/1p4mSIse4PEPx9pDR7TMDgL6nPVnLpzFi?usp=share_link> (Accessed 07 March 23).
33. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, arXiv:1512.03385
34. Andriyanov N.A., Dementiev V.E., Tashlinskiy A.G. Deep Markov models of multidimensional random fields, Procedia Computer Science, 2020, Vol. 176, pp. 1289–1298, DOI: 10.1016/j.procs.2020.09.138
35. Andriyanov N., Khasanshin I., Utkin D., et al. Intelligent system for estimation of the spatial position of apples based on YOLOv3 and Real Sense depth camera D415, Symmetry, 2022, Vol. 14, No. 1, article 148, DOI: 10.3390/sym14010148
36. Kuznetsova A., Maleva T., Soloviev V. Ispolzovanie algorythma YOLOv3 c procedurami pred- i postobrabotki dlya obnaruzheniya plodov robotom dlya sbora yablok [Using the YOLOv3 algorithm with preand post-processing procedures for fruit detection by an apple-picking robot], Prikladnaya informatika [Journal of Applied Informatics], 2020, Vol. 15, No. 4, pp. 64-74, DOI: 10.37791/2687-0649-2020-15-4-64-74
37. Kalashnikov V., Soloviev V. Prilozheniya computernogo zreniya v gornodobyvayushchey promyshlennosti [Applications of computer vision in the mining industry], Prikladnaya informatika [Journal of Applied Informatics], 2023, Vol. 18, No. 1, pp.4-21, DOI: 10.37791/2687-0649-2023-18-1-4-21

**Информация об авторах**

*Ахмад Авс*,аспирант кафедры «робототехнические системы и мехатроника» МГТУ им. Н. Э. Баумана, руководитель группы мехатроники и робототехники ООО «ЦИАРС»,ассистент Департамента анализа данных и машинного обучения Финансового университета, *email:* aws.ahmad318@ gmail.com

*Андриянов Никита Андреевич,* доцент Департамента анализа данных и машинного обучения Финансового университета, *email:* [NAAndriyanov@fa.ru](mailto:NAAndriyanov@fa.ru)

*Соловьев Владимир Игоревич*,генеральный директор ООО «ЦИАРС»,заведующий кафедрой «Прикладной искусственный интеллект» МТУСИ, *email:* vs@ciars.ai

*Соломатин Денис Алексеевич*, бакалавр Департамента анализа данных и машинного обучения Финансового университета, *email:* deniso2001@gmail.com

**Author information**

*Ahmad Aws*, PhD student of the of Robotic Systems and Mechatronics Department, Bauman Moscow State Technical University; Team Leader, Mechatronics and Robotics, CIARS Llc; Assistant Professor, Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia, *email:* aws.ahmad318@gmail.com

*Nikita Andriyanov*, Associate Professor, Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia, *email:* [NAAndriyanov@fa.ru](mailto:NAAndriyanov@fa.ru)

*Vladimir Soloviev*, Dr. Sci. (Econ.), CEO, CIARS Llc; Professor and Chair, Department of Applied Artificial Intelligence, Moscow Technical University of Communications and Informatics, Moscow, Russia, vs@ciars.ai

*Denis Solomatin*, Bachelor, Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia, *email:* deniso2001@gm