

Информатика и вычислительная техника

Informatics and computer engineering

Научная статья

УДК 004.896

DOI: 10.14529/ctcr230201

ПРИМЕНЕНИЕ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ АУГМЕНТАЦИИ И ГЕНЕРАЦИИ ПОДВОДНОГО НАБОРА ДАННЫХ С ПРОМЫШЛЕННЫМИ ОБЪЕКТАМИ

А. Ахмад^{1, 2, 3}, aws.ahmad318@gmail.com

Н.А. Андриянов³, naandriyanov@fa.ru, <https://orcid.org/0000-0003-0735-7697>

В.И. Соловьев^{2, 4}, vs@ciars.ai, <https://orcid.org/0000-0003-0338-1227>

Д.А. Соломатин³, deniso2001@gmail.com

¹ Московский государственный технический университет имени Н.Э. Баумана, Москва, Россия

² ООО «ЦИАРС», Москва, Россия

³ Финансовый университет при Правительстве Российской Федерации, Москва, Россия

⁴ Московский технический университет связи и информатики, Москва, Россия

Аннотация. Цель исследования: разработка метода глубокого обучения для аугментации и генерации проблемно-ориентированного набора данных, содержащего промышленные объекты, в том числе разработка более эффективных алгоритмов генерации и аугментации данных, основанных на глубоком обучении, которые позволяют создавать более разнообразные и реалистичные данные, соответствующие промышленным объектам, которые могут быть перенесены из одного стилевого домена в другой. Поставленная в исследовании цель связана с наличием актуальной научно-технической задачи обеспечения компьютерного зрения в системах, работающих в подводной среде. Это могут быть автономные необитаемые подводные аппараты, которые ищут прорывы в трубопроводах, анализируют утечку нефти, движение косяков рыб и т. п. Однако недостаточного количества данных, содержащих описанные объекты в условиях их реального существования, сегодня нет. Таким образом, необходимо обеспечить обучающую выборку реалистичными изображениями. **Методы исследования:** архитектура CycleGAN, обеспечивающая преобразование набора данных, содержащего изображения различных объектов, сделанные в лабораторной или в обычной наземной среде, в набор данных, содержащий характеристики подводной среды. Для оценки разработанного алгоритма аугментации предлагается использовать классификацию изображений по доменам, которая может быть выполнена с помощью сверточной нейронной сети ResNet. **Результаты исследования.** Представлен инструмент для решения проблемы отсутствия подводных наборов данных, разработана модель глубокого обучения, которая применяется для создания изображений с подводными элементами. Модель работает по принципу циклической генеративно-состязательной сети, которая получает на вход реальное изображение промышленного объекта в надводных условиях, а на выход возвращает сгенерированное изображение того же промышленного объекта в подводных условиях. Визуальный анализ изображений показывает, что такой метод достаточно адекватен. Кроме того, проверка на классификационной модели показала почти 100%-ную способность нейросети различать домены. **Заключение.** Исследование показало, что модель CycleGAN можно использовать для создания изображений различных объектов в подводной среде. В будущем возможен поиск дополнительных процедур аугментации, кроме того, могут быть использованы аугментации сгенерированного набора изображений, что также обеспечит исследователей и разработчиков достаточным материалом с промышленными объектами в подводной среде. Это может повысить качество разработок.

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

Для цитирования: Применение глубокого обучения для аугментации и генерации подводного набора данных с промышленными объектами / А. Ахмад, Н.А. Андриянов, В.И. Соловьев, Д.А. Соломатин // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». 2023. Т. 23, № 2. С. 5–16. DOI: 10.14529/ctcr230201

APPLICATION OF DEEP LEARNING FOR AUGMENTATION AND GENERATION OF AN UNDERWATER DATA SET WITH INDUSTRIAL FACILITIES

A. Ahmad^{1, 2, 3}, aws.ahmad318@gmail.com

N.A. Andriyanov³, naandriyanov@fa.ru, <https://orcid.org/0000-0003-0735-7697>

V.I. Soloviev^{2, 4}, vs@ciars.ai, <https://orcid.org/0000-0003-0338-1227>

D.A. Solomatin³, deniso2001@gmail.com

¹ Bauman Moscow State Technical University, Moscow, Russia

² CIARS LLC, Moscow, Russia

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

⁴ Moscow 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 dataset containing industrial objects, including the development of more efficient data generation and augmentation algorithms based on deep learning, which allow you to create more diverse and realistic data corresponding to industrial objects that can be transferred from one style domain to another. The goal set in the study is related to the actual scientific and technical problem of providing computer vision in systems operating in the underwater environment. These can be autonomous uninhabited underwater vehicles that look for breaks in pipelines, analyze oil leaks, the movement of schools of fish, etc. However, today there is not enough data containing the described objects in the conditions of their real existence. Thus, it is necessary to provide the training sample with realistic images.

Research methods: the CycleGAN architecture, which converts a dataset containing images of various objects taken in a laboratory or in a conventional aboveground environment into a dataset containing the characteristics of an underwater environment. To evaluate the developed augmentation algorithm, it is proposed to use image classification by domains, which can be performed using the ResNet convolutional neural network. **Results of the study.** A tool is presented to solve the problem of the lack of underwater datasets, a deep learning model is developed, which is used to create images with underwater elements. The model works on the principle of a cyclic generative adversarial network, which receives a real image of an industrial facility in surface conditions as an input, and returns a generated image of the same industrial facility in underwater conditions as an output. Visual analysis of images shows that this method is quite adequate. In addition, a test on the classification model showed almost 100% ability of the neural network to distinguish between domains. **Conclusion.** The study showed that the CycleGAN model can be used to create images of various objects in the underwater environment. In the future, it is possible to search for additional augmentation procedures, in addition, augmentations of the generated set can be used images, which will also provide researchers and developers with sufficient material with industrial facilities in the underwater environment. This can improve the quality of developments.

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

For citation: Ahmad A., Andriyanov N.A., Soloviev V.I., Solomatin D.A. Application of deep learning for augmentation and generation of an underwater data set with industrial facilities. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*. 2023;23(2):5–16. (In Russ.) DOI: 10.14529/ctcr230201

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 подводных 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 изображениях. 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 итоге 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 водолозов 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.

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. Неклас-

сические методы аугментации включают внесение изменений в набор данных путем применения более продвинутых методов, таких как добавление шума, обрезка и изменение цвета изображений [26].

Наборы изображений подводной среды
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

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-сопоставительные 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-сопоставительные сети в основном используются для создания синтетических изображений, которые следуют тому же распределению вероятностей, что и реальные изображения. CycleGAN – это хорошо известная архитектура генеративно-сопоставительных 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.

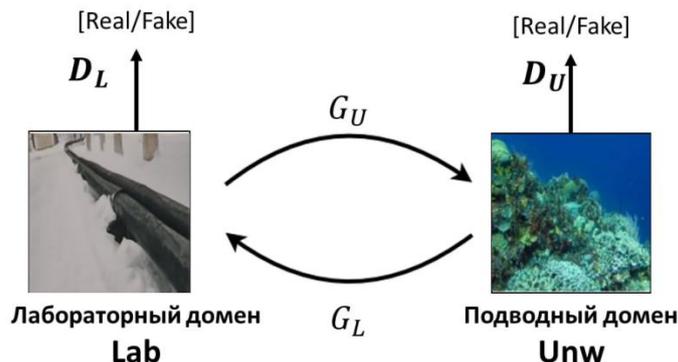


Рис. 1. Схема работы CycleGAN
Fig. 1. CycleGAN scheme

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

$$\begin{aligned} G_L &: Unw \rightarrow Lab; \\ G_U &: Lab \rightarrow Unw. \end{aligned} \quad (1)$$

G_L и G_U – это генераторы, которые берут изображение из одного домена и переводят его в другой. G_L отображает картинку из Unw в Lab ($x \rightarrow y$), тогда как G_U идет в противоположном направлении, отображая из Lab в Unw ($y \rightarrow x$).

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

$$\begin{aligned} D_L &: \text{отличает } x \text{ от } G_L(y); \\ D_U &: \text{отличает } y \text{ от } G_U(x). \end{aligned} \quad (2)$$

D_L дискриминатор обеспечивает состязательное обучение (adversarial learning) для G_L , а D_U делает то же самое для G_U .

Оба генератора G_L и G_U 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:

$$\begin{aligned} Loss_{adv}(G_U, D_U, Lab) &= \frac{1}{m} \sum_1^m [1 - D_U(G_U(x_i))]^2; \\ Loss_{adv}(G_L, D_L, Unw) &= \frac{1}{m} \sum_1^m [1 - D_L(G_L(y_i))]^2, \end{aligned} \quad (3)$$

где $x_i \in Lab$, $y_i \in Unw$;

2) Cycle Consistency Losses – потери согласованности цикла для предотвращения противоречия изученных отображений G_L и G_U друг другу:

$$Loss_{cy}(G_L, G_U, Lab, Unw) = \frac{1}{m} \sum_1^m [G_L(G_U(x_i) - x_i)] + [G_U(G_L(y_i) - y_i)]. \quad (4)$$

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 :

$$Loss_{Full} = Loss_{adv} + \lambda Loss_{cy}. \quad (5)$$

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

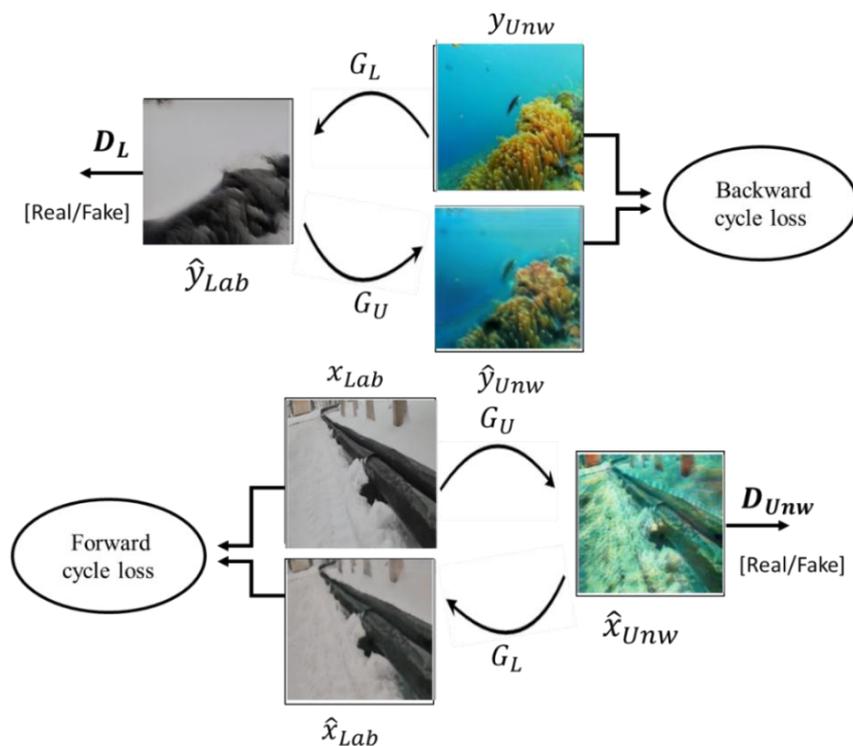


Рис. 2. Потери согласованности цикла
Fig. 2. Cycle consistency losses

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].

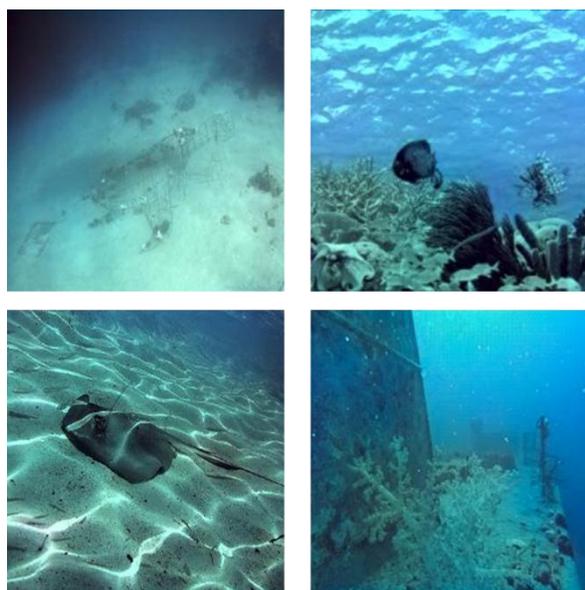


Рис. 3. Набор изображений подводной среды
(underwater_ImageNet) [19]
Fig. 3. Underwater environment image set
(underwater_ImageNet) [19]



Рис. 4. Базовый набор изображений [32]
Fig. 4. The basic image set [32]

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.

Архитектура модели CycleGAN

Архитектура модели CycleGAN представлена на рис 5. В этой модели характеристики изображения подводного домена y_{Unw} преобразуются в лабораторный домен \hat{y}_{Lab} с помощью генератора G_L .

Затем генератор G_U преобразует изображения лабораторного домена x_{Lab} в подводный домен, что в итоге приводит к созданию искусственных изображений \hat{x}_{Unw} .

Дискриминаторы D_L, D_{Unw} отвечают за мониторинг и сравнение искусственных изображений $\hat{y}_{Lab}, \hat{x}_{Unw}$ в каждом домене с подлинными на самой последней фазе сети, при этом вычисляются состязательные подери $Loss_{adv}$ (3), по которым оптимизируется окончательный вывод изображений дискриминаторами.

Затем на каждом этапе процесса модель реконструирует искусственные изображения $\hat{y}_{Lab}, \hat{x}_{Unw}$ в исходный домен $\hat{x}_{Lab}, \hat{y}_{Unw}$ и рассчитывает потери согласованности цикла $Loss_{cy}$ (4).

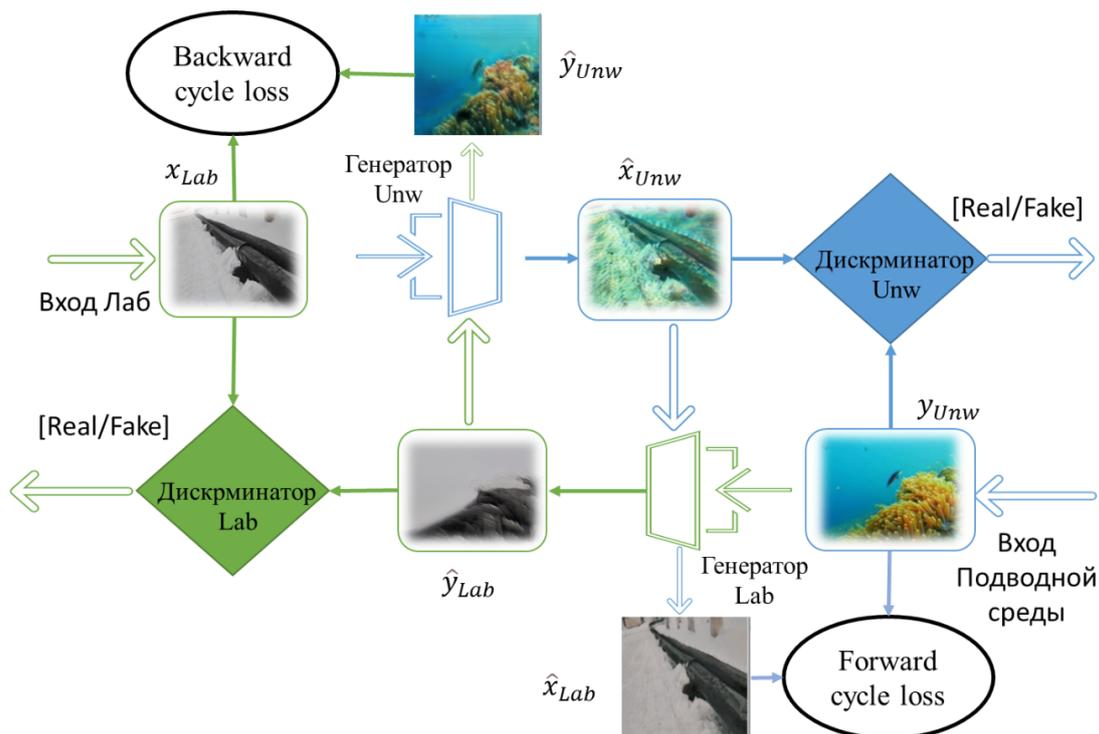


Рис. 5. Общая схема CycleGAN
Fig. 5. The general scheme of CycleGAN

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 FC и установлен выход на 2 класса. Был обучен только последний линейный слой FC с использованием 3000 изображений подводной среды (из набора Underwater_ImageNet) и 500 изображений из базового набора.

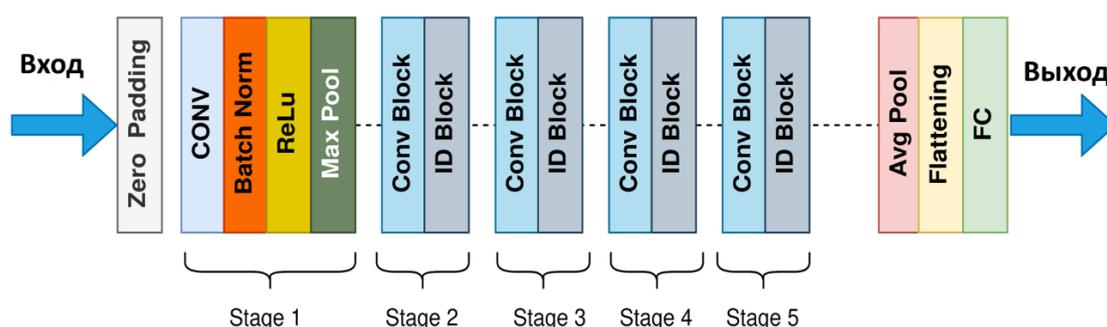


Рис. 7. Архитектура ResNet-50
Fig. 7. ResNet-50 architecture

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

Матрица ошибок классификации, представленная на рис. 9, позволяет оценить точность классификации $Precision = TP / (TP + FP) = 0,99865$, полноту $Recall = TP / (TP + FN) = 1,00000$ и метрику $F_1 = 2Precision \cdot Recall / (Precision + Recall) = 0,99933$.

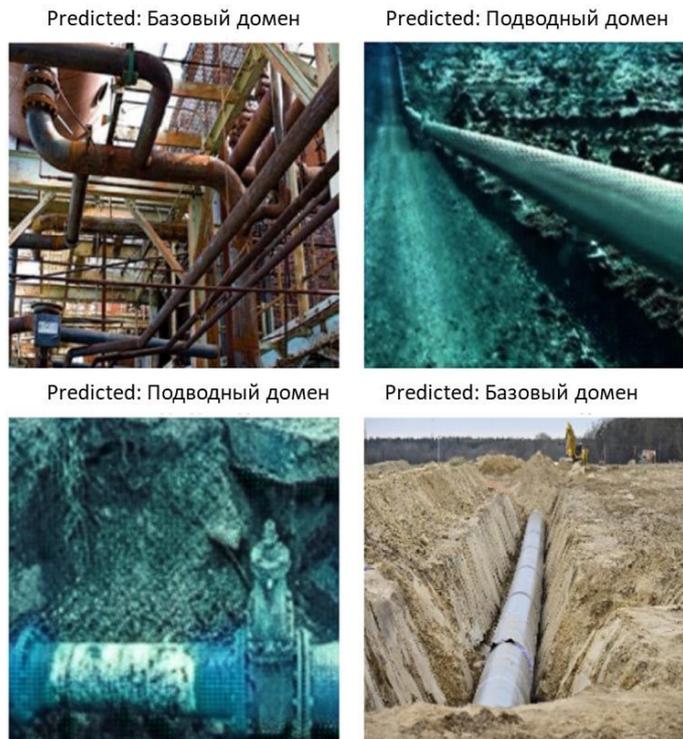


Рис. 8. Примеры классификации изображений
Fig. 8. Image classification examples



Рис. 9. Матрица ошибок классификации
Fig. 9. 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 ($F_1 = 0,99933$ на тестовой выборке из $N = 805$ изображений).

Сгенерированный новый набор изображений можно использовать для решения задач сегментации изображений и обнаружения объектов с помощью трансферного обучения [34–38].

Список литературы/References

1. Ridolfi A., Conti R., Costanzi R. A dynamic manipulation strategy for an intervention: Autonomous underwater vehicle. *Advances in Robotics and Automation*. 2015;4(2):100–132. DOI: 10.4172/2168-9695.1000132
2. Wynn R.B., Huvenne V.A.I., Le Bas T.P. Autonomous underwater vehicles (AUVs): Their past, present and future contributions to the advancement of marine geosciences. *Marine Geology*. 2014;352:451–468. 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;88:627–630. DOI: 10.1016/j.oceaneng.2014.04.027
4. Manzanilla A., Reyes S., Garcia M. Autonomous navigation for unmanned underwater vehicles: Real-time experiments using computer vision. *IEEE Robotics and Automation Letters*. 2019;4(2):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;13(1):1–16. 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;46(1):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. In: *OCEANS 2017 – Aberdeen*. Aberdeen, UK; 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;6:17880–17886. 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;27(9):1837–1844. DOI: 10.1007/s00542-019-04694-8
10. Girshick R. Fast R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Araucano Park; 2015. Vol. 1. P. 1440–1448.
11. Bochkovskiy A., Wang C.Y., Liao H.Y.M. YOLOv4: Optimal speed and accuracy of object detection. Available at: <https://arxiv.org/abs/2004.10934> (accessed 03.04.2023).
12. Cazzato D. A survey of computer vision methods for 2d object detection from unmanned aerial vehicles. *Journal of Imaging*. 2020;6:1–8. DOI: 10.3390/jimaging6080078
13. Dakhil R.A., Khayyat A.R.H. Review on deep learning technique for underwater object detection. In: *Proceedings of the 3rd International Conference on Data Science and Machine Learning (DSML)*. Copenhagen; 2022. Vol. 1. P. 49–63.
14. Jian M., Qi Q., Yu H. The extended marine underwater environment database and baseline evaluations. *Applied Soft Computing*. 2019;80:425–437. DOI: 10.1016/j.asoc.2019.04.025
15. Liu R., Hou M., Fan X. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light. *IEEE Transactions on Circuits and Systems for Video Technology*. 2020;30(12):4861–4875. DOI: 10.1109/TCSVT.2019.2963772
16. Islam M.J., Edge C., Xiao Y. Semantic segmentation of underwater imagery: Dataset and benchmark. In: *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Las Vegas; 2020. Vol. 1. P. 1–6. DOI: 10.1109/IROS45743.2020.9340821
17. Hong J., Fulton M., Sattar J. TrashCan: A semantically-segmented dataset towards visual detection of marine debris. Available at: <https://arxiv.org/abs/2007.08097> (accessed 01.04.2023).
18. Fabbri C., Islam M.J., Sattar J. Enhancing underwater imagery using generative adversarial networks. In: *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane; 2018. Vol. 1. P. 7159–7165.

19. Underwater_ImageNet Dataset. Available at: <https://drive.google.com/file/d/1LOM-2A1BSLaFjCY2EEK3DA2Lo37rNw-7/view> (accessed 07.03.23).
20. Test-lhnog Dataset. Available at: <https://universe.roboflow.com/new-workspace-qnaXk/test-lhnog> (accessed 07.03.23).
21. Human Dataset. Available at: <https://universe.roboflow.com/school-9j0uz/human-ighgb> (accessed 07.03.23).
22. Please Dataset. Available at: <https://universe.roboflow.com/new-workspace-bpkjt/please> (accessed 07.03.23).
23. Pipes Dataset. Available at: <https://universe.roboflow.com/pipes-nty6m/p-w1typ> (accessed 07.03.23).
24. Perez L., Wang J. The effectiveness of data augmentation in image classification using deep learning. Available at: <https://arxiv.org/abs/1712.04621> (accessed 24.03.2023).
25. Andriyanov N.A., Andriyanov D.A. The using of data augmentation in machine learning in image processing tasks in the face of datascarcity. *Journal of Physics Conference Series*. 2020;1661(1):012018. DOI: 10.1088/1742-6596/1661/1/012018
26. Buslaev A., Igloukov V.I., Khvedchenya E., Parinov A., Druzhinin M., Kalinin A.A. Albumentations: Fast and flexible image augmentations. *Information*. 2020;11(2):1–20. DOI: 10.3390/info11020125
27. Goodfellow I.J., Pouget-Abadie J., Mirza M. Generative adversarial networks. *Communications of the ACM*. 2020;63(1):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. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Venice; 2017. Vol. 1. P. 2223–2232.
29. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. Available at: <https://arxiv.org/abs/1505.04597> (accessed 02.04.2023).
30. Polymenis I., Haroutunian M., Norman R., Trodden D. Virtual underwater datasets for autonomous inspections. *Journal of Marine Science and Engineering*. 2022;9:1–18. 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. Available at: <https://arxiv.org/abs/1806.07777> (accessed 03.04.2023).
32. Набор изображений механического и промышленного оборудования [*Nabor izobrazheniy mekhanicheskogo i promyshlennogo oborudovaniya* [Dataset of images of mechanical and industrial equipment]]. Available at: https://drive.google.com/drive/folders/1p4mSIse4PEPx9pDR7TMDgL6nPVnLpzFi?usp=share_link (accessed 25.03.23)].
33. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas; 2016. Vol. 1. P. 770–778.
34. Andriyanov N.A., Dementiev V.E., Tashlinskiy A.G. Deep Markov models of multidimensional random fields. *Procedia Computer Science*. 2020;176:1289–1298. DOI: 10.1016/j.procs.2020.09.138
35. Andriyanov N., Khasanshin I., Utkin D., Gataullin T., Ignar S., Shumaev V., Soloviev V. Intelligent system for estimation of the spatial position of apples based on YOLOv3 and Real Sense depth camera D415. *Symmetry*. 2022;14(1):148. DOI: 10.3390/sym14010148
36. Kuznetsova A.A., Maleva T.V., Soloviev V.I. Using the YOLOv3 algorithm with pre- and post-processing procedures for fruit detection by an apple-picking robot. *Journal of Applied Informatics*. 2020;15(4):64–74. (In Russ.) DOI: 10.37791/2687-0649-2020-15-4-64-74
37. Kalashnikov V., Soloviev V. Applications of computer vision in the mining industry. *Journal of Applied Informatics*. 2023;8(1):4–21. (In Russ.) DOI: 10.37791/2687-0649-2023-18-1-4-21
38. Andriyanov N.A., Volkov A.I., Volkov A.N., Gladkikh A.A. Research of recognition accuracy of dangerous and safe x-ray baggage images using neural network transfer learning. *IOP Conference Series: Materials Science and Engineering*. 2021;1061:012002. DOI: 10.1088/1757-899X/1061/1/012002

Информация об авторах

Ахмад Авс, аспирант кафедры робототехнических систем и мехатроники, Московский государственный технический университет имени Н.Э. Баумана, Москва, Россия; руководитель группы мехатроники и робототехники, ООО «ЦИАРС», Москва, Россия; ассистент Департамента анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации, Москва, Россия; aws.ahmad318@gmail.com.

Андрянов Никита Андреевич, канд. техн. наук, доц. Департамента анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации, Москва, Россия; naandriyanov@fa.ru.

Соловьев Владимир Игоревич, д-р экон. наук, проф., генеральный директор, ООО «ЦИАРС», Москва, Россия; заведующий кафедрой прикладного искусственного интеллекта, Московский технический университет связи и информатики, Москва, Россия; vs@ciars.ai.

Соломатин Денис Алексеевич, бакалавр Департамента анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации, Москва, Россия; deniso2001@gmail.com.

Information about the authors

Aws Ahmad, Postgraduate Student of the Department of Robotic Systems and Mechatronics, Bauman Moscow State Technical University, Moscow, Russia; Head of Mechatronics and Robotics Group, CIARS LLC, Moscow, Russia; Assistant of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia; aws.ahmad318@gmail.com.

Nikita A. Andriyanov, Cand. Sci. (Eng.), Ass. Prof. of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia; naandriyanov@fa.ru.

Vladimir I. Soloviev, Dr. Sci. (Econ.), Prof., CEO, CIARS LLC, Moscow, Russia; Head of the Department of Applied Artificial Intelligence, Moscow Technical University of Communications and Informatics, Moscow, Russia; vs@ciars.ai.

Denis A. Solomatin, Bachelor of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, Russia; deniso2001@gmail.com.

Статья поступила в редакцию 11.03.2023

The article was submitted 11.03.2023