

Автоматизированный поиск повреждений подводных сетчатых ограждений

С. И. СМАГИН¹, А. С. СМАГИН^{2,*}

¹Вычислительный центр ДВО РАН, 680000, Хабаровск, Россия

²Хабаровский федеральный исследовательский центр ДВО РАН, 680000, Хабаровск, Россия

*Контактный автор: Смагин Алексей Сергеевич, e-mail: smaginkhv@gmail.com

Поступила 04 апреля 2023 г., доработана 05 мая 2023 г., принята в печать 12 мая 2023 г.

Рассматриваются вопросы автоматизации визуального мониторинга состояния сетчатых ограждающих конструкций, которые применяются на морских рыболовных фермах. Разработаны алгоритмы и комплекс компьютерных программ на языке Python для поиска повреждений подводных сетчатых ограждений по их фото и видеоизображениям методами компьютерного зрения и машинного обучения. Приведены результаты работы программного комплекса. Показано, что технологии компьютерного зрения и машинного обучения эффективно справляются с контролем целостности ячеек подводного сетчатого ограждения.

Ключевые слова: автоматизация, мониторинг, сетчатое ограждение, алгоритм, компьютерное зрение, машинное обучение, подводный аппарат.

Цитирование: Смагин С.И., Смагин А.С. Автоматизированный поиск повреждений подводных сетчатых ограждений. Вычислительные технологии. 2023; 28(3):182–194. DOI:10.25743/ICT.2023.28.3.011.

Введение

Рыбоводные фермы чаще всего представляет собой погруженные в воду клетки (садки), основа которых обычно состоит из металлического каркаса, обтянутого сеткой из синтетического материала. Самые распространенные материалы сетки — это полиамид и полиэфир. Условия эксплуатации подводных рыболовных ферм накладывают жесткие ограничения на непосредственное участие человека в их мониторинге и обслуживании и стимулируют применение и развитие безлюдных технологий производства морских биоресурсов.

Применяются два основных способа мониторинга ограждений подводных аквакультурных ферм: при помощи водолазов и с использованием телеуправляемых необитаемых подводных аппаратов (ТНПА), которые оснащены специальным оборудованием. Аппарат погружается в воду и управляется из кабины на поверхности воды оператором или группой операторов (пилотом, навигатором и др.). Работа с таким аппаратом намного дешевле дорогостоящих водолазных работ. Однако для эффективного применения ТНПА необходимы квалифицированные операторы, их обучение и работа требуют серьезных финансовых вложений и больших временных затрат [1].

В ближайшей перспективе получит развитие еще один способ мониторинга подводных объектов, который основан на применении автономных необитаемых подводных аппаратов (АНПА). Однако современные АНПА пока не способны выполнять такой же

широкий спектр задач, как водолазы или даже ТНПА. На предприятиях аквакультуры их можно использовать для решения относительно простых задач, таких как поиск объектов под водой, фото- и видеосъемка сетчатых ограждений, а это весьма значительная часть подводных работ [2].

Под водой спутниковая связь не работает, поэтому для коммуникации подводных роботов с людьми используется гидроакустическая связь. У такой связи есть ограничения по дальности и скорости передачи данных, поэтому для интенсивного обмена данными АНПА обычно всплывают на поверхность воды. При эксплуатации АНПА требуются лишь своевременное техническое обслуживание и контроль их действий в сложных ситуациях [3].

Алгоритмы системы компьютерного зрения АНПА должны корректно работать в разнообразных и сложных условиях реальной водной среды и надежно выполнять задачи мониторинга, в том числе определять повреждения и загрязнения подводных сетчатых ограждений. Кроме того, весьма желательно обеспечить автономную обработку данных на борту АНПА, чтобы иметь возможность прервать миссию в случае обнаружения критического повреждения и сократить время между обнаружением повреждения и передачей сигнала технической службе о его местоположении. Автономность обработки данных и определения нужных характеристик объектов достигается применением технологий компьютерного зрения и машинного обучения [4, 5].

В настоящей статье представлены результаты разработки и исследования алгоритмов обработки фото- и видеоизображений, предназначенных для решения задач автоматизации визуального мониторинга сетчатых ограждающих конструкций для морских рыбоводных ферм [6]. Предложенные алгоритмы компьютерного зрения и машинного обучения позволяют выявлять наличие порывов в сетчатых ограждениях по фотоизображениям, полученным в подводных условиях. Они реализованы в виде комплекса программ на языке Python с использованием библиотеки компьютерного зрения OpenCV и модифицированной архитектуры сверточной нейронной сети (СНС) U-Net. Тестирование программного комплекса на реальных данных, полученных с борта ТНПА, показало возможность его использования в системах компьютерного зрения подводных роботов.

1. Сегментация изображений

Сегментация — одна из основных задач обработки и анализа изображений. Она заключается в разделении изображения на области, для которых выполняются определенные критерии однородности (например, яркости). Для сегментации изображений при помощи “классических” методов компьютерного зрения характерно то, что в их основе лежит информация о пространственных признаках объекта и в них отсутствует нейросетевая составляющая.

Основным методом сегментации в компьютерном зрении является пороговая обработка изображения. Перед ее выполнением целесообразно предварительно произвести улучшение изображения. Для этого, а именно для компенсации артефактов, связанных с освещением, в реализованном программном комплексе для поиска повреждений подводных сетчатых ограждений по их фото- и видеоизображениям применяется алгоритм multi scale retinex with color restoration (MSRCR) [5].

Изображения объектов, полученные фото- и видеокамерами подводных роботов в реальных условиях их применения, отличаются от снимков, сделанных без погружения

в подводную среду. Чаще всего эти отличия проявляются в более высоком уровне оптических искажений, слабой освещенности и контрастности подводных фото- и видеоизображений.

Основная сложность получения качественных снимков при решении задач мониторинга подводных объектов обусловлена поглощением и рассеянием света как самой водой, так и растворенными в ней органическими веществами и малыми взвешенными частицами. Оптические характеристики водной среды, которая поглощает и рассеивает свет, оказывают значительное влияние на дальность наблюдения, контрастность и размытость границ получаемых изображений. Недостаток или полное отсутствие естественной освещенности на рабочих глубинах компенсируются оснащением системы подводного видения робота внешней подсветкой. Такая подсветка обеспечивает повышение уровня освещенности объекта наблюдения на снимке, но также усиливает неравномерность распределения яркости на изображении.

Использование глобальных порогов бинаризации при обработке снимков подводных объектов не всегда дает корректные результаты. На одних участках изображения автоматически определяемый порог бинаризации позволяет получить достаточно хорошо визуализируемые объекты, тогда как на других необходимые для анализа объекты становятся “засвеченными” или, наоборот, “затененными”.

Результаты применения классических методов компьютерного зрения, использующих пороговую бинаризацию, показали, что они успешно справляются с сегментацией сетчатого ограждения на не сильно зашумленных изображениях. Однако этого недостаточно для полного отсека фона и посторонних объектов на фотоснимках, полученных в реальных подводных условиях. Для хорошей работы этих алгоритмов требуется тонкая настройка для каждого изображения в отдельности. На реальных подводных изображениях сетчатого ограждения уровень освещенности сильно меняется в зависимости от глубины в момент съемки. Поэтому настройка параметров алгоритма для одного снимка не дает приемлемых результатов для снимков, сделанных в других условиях.

При создании комплекса программ для поиска повреждений сетчатых ограждений был рассмотрен и реализован метод семантической сегментации подводных изображений сетчатых ограждений [7]. Цель метода — точная классификация и семантическая сегментация каждого объекта, представленного на изображении. Этот вид сегментации изображений подразумевает отнесение каждого пикселя к определенному классу (объекту). Семантическая сегментация заметно отличается от обычной сегментации, когда области на изображении объединяются по принципу цветового или текстурного сходства. Объекты могут состоять из существенно различающихся по фотометрическим характеристикам элементов и иметь значительный разброс показателей внутри одного класса [8].

Решение задачи семантической сегментации включает следующие этапы: предварительная обработка данных, создание обучающих выборок, выбор функции ошибок (потерь), выбор или разработка архитектуры нейронной сети, обучение нейронной сети и выполнение сегментации, постобработка полученных карт разметки.

Для оценки точности сегментации была выбрана метрика IoU (intersection over union) [9]. Она определяется как отношение площади пересечения истинной и предсказанной областей сегментации к площади их объединения. Если она рассчитывается для нескольких классов, то сначала вычисляется IoU для каждого класса, а затем берется их среднее значение.

Стоит отметить, что стандартные метрики не в полном объеме отражают качество работы алгоритмов для сегментации сетчатого ограждения. В нашем случае важно, чтобы можно было по маске производить оценку состояния сетчатого ограждения. А значит, уместнее сравнивать количество целых ячеек ограждения на исходной маске с числом целых ячеек на маске, сгенерированной СНС.

В глубоком машинном обучении задача подготовки качественного набора исходных данных (обучающей выборки, датасета) для обучения нейронной сети — одна из основных. Главная проблема подготовки хорошей обучающей выборки состоит в сложности поиска нужных изображений, так как искомые объекты на разных изображениях должны быть похожими. Желательно, чтобы обучающая выборка состояла из большого числа изображений. В разных источниках рекомендуемый размер обучающей выборки составляет тысячу изображений и более. Но таким большим числом подходящих изображений исследователи не располагали, выборка состояла из 80 подводных снимков [10–12]. Эти фотоизображения были предоставлены коллегами из Китайского университета океанологии (г. Циндао). Съемка производилась с борта ТНПА на местной рыбоводной ферме.

Для расширения (аугментации) данной выборки использованы различные фильтры (*sharpness*, *blur*), изменяющие качество изображения. Под аугментацией понимается увеличение выборки данных для обучения через модификацию существующих данных. Также использовались поворот изображения на угол от 5 до 20 град. и зеркальное отражение с последующим поворотом на малый угол [10, 13]. Повороты применялись таким образом, чтобы в обучающей выборке не было похожих друг на друга изображений. При формировании выборки учитывалось, что подводные изображения сильно отличаются по яркости и равномерности освещения, размытости, контрастности, наличию сторонних объектов. В отдельных случаях это делает их мало различимыми даже для человека.

Таким образом, выборку из 80 изначально размеченных изображений удалось расширить до 7680. Разметка производилась с помощью фоторедактора GIMP. Ввиду сложности структуры объекта разметка одного изображения занимала около 15 мин. Разметки аугментировались аналогично соответствующим оригинальным изображениям. Примеры изображений из выборки и их разметки представлены на рис. 1. В верхнем



Рис. 1. Примеры изображений и их разметки

Fig. 1. Examples of images and their markup

ряду первые два изображения — исходные, третье и четвертое — аугментированные, соответствующие разметки приведены в нижнем ряду. Итоговая выборка разбита на обучающую (4608 изображений) и тестовую (3072 изображения).

2. Выбор архитектуры нейронной сети для сегментации

Для задачи семантической сегментации предложено много разнообразных методов решения. Однако результаты их сравнения на открытых наборах данных, например ISPRS Semantic Labeling Contest [14], показывают значительное превосходство алгоритмов, основанных на СНС в комбинации с различными подходами к пред- и постобработке изображений.

Нами проведено сравнение результатов сегментации для базовых архитектур СНС U-Net [15], Seg-Net [16] и Mask R-CNN [17]. Лучшие результаты сегментации показала СНС на архитектуре U-Net. Обученная на расширенном наборе данных, она показала точность семантической сегментации в 1.5 раза выше, чем при обучении на базовой выборке, состоящей из исходных изображений.

Графики значений функции ошибки для тренировочных (train loss) и тестовых (val loss) данных и результаты семантической сегментации представлены на рис. 2 и 3. На рис. 2 видно, что после 30-й эпохи происходит переобучение, поэтому на данном этапе процесс обучения был завершен.

В общей сложности протестировано одиннадцать разных архитектур СНС. Шесть из них представляют собой модификации архитектуры U-Net. Основных модификаций можно выделить четыре. Это U-Net с обучаемым upsampling, энкодером VGG16BN, residual-блоками и энкодером ResNet34. Чтобы исключить проблему переобучения и повысить точность сегментации выбрана функция ошибки Boundary loss [18]. Результаты проверки точности сегментации по метрике IoU для всех протестированных архитектур сверточных нейронных сетей (СНС) приведены в табл. 1. Следуя авторам работ [19–21], будем считать, что значения точности, превышающие 50 %, являются хорошими показателями по этой метрике.

Модель на архитектуре U-Net с residual-блоками показала точность семантической сегментации по метрике IoU, равную 62.8 %. Этот показатель на 1 % ниже, чем для модели на архитектуре U-Net, и на 2.5 % ниже, чем для модели на архитектуре U-Net

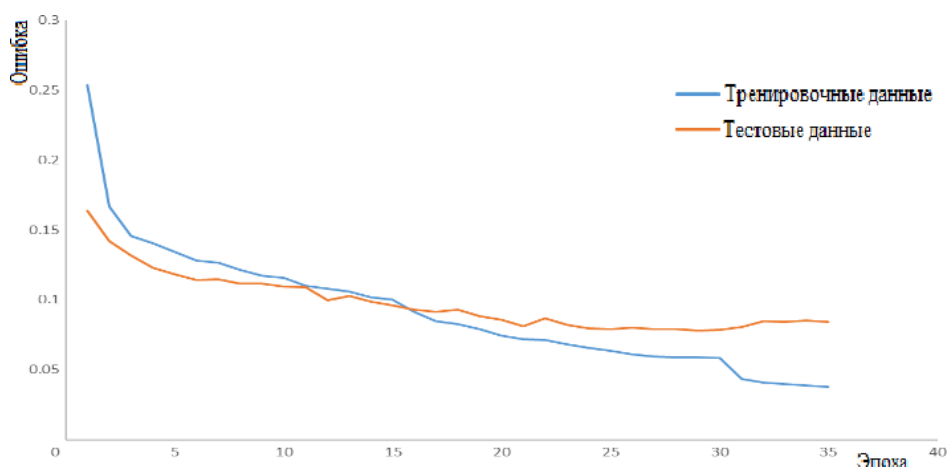


Рис. 2. Графики значений функции ошибки для U-Net

Fig. 2. Graphs of error function values for U-Net

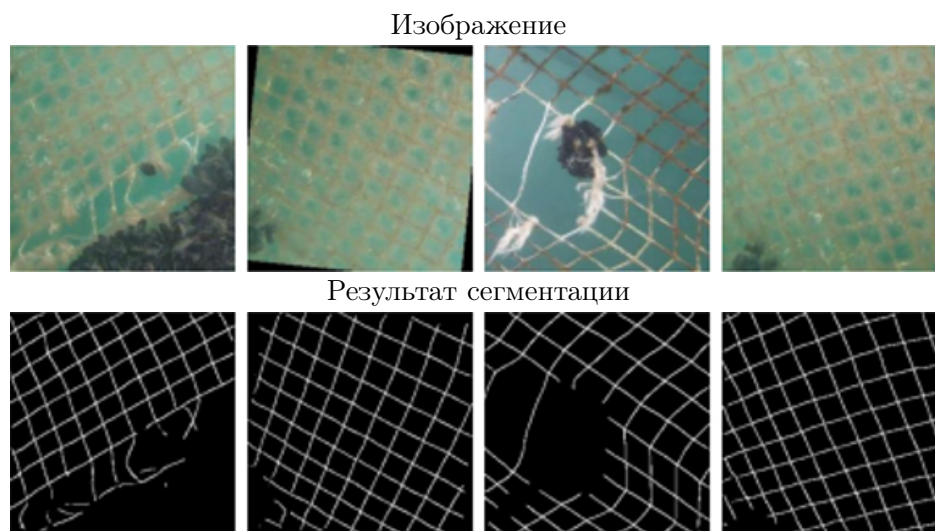


Рис. 3. Результаты сегментации с помощью нейросети на архитектуре U-Net
 Fig. 3. Segmentation results using a neural network on the U-Net architecture

Т а б л и ц а 1. Точность сегментации для различных архитектур СНС на тестовой выборке
 Table 1. Segmentation accuracy for various CNN architectures

Архитектура	Точность маски (IoU), %
U-Net с энкодером VGG16BN	65.3
U-Net	63.8
U-Net с residual-блоками	62.8
U-Net с обучаемым upsampling	62.4
Fast FCN	62.3
U-Net с энкодером ResNet34 12 blocks	61.2
U-Net с энкодером ResNet50	60.8
U-Net с энкодером ResNet34	60.2
MobileNetV2	56.6
SegNet	55.0
Mask R-CNN	21.0

с энкодером VGG16BN. Полученный результат требует дальнейшего исследования. Не до конца ясно, почему модель на архитектуре U-Net с residual-блоками уступила по точности моделям на архитектурах U-Net и U-Net с энкодером VGG16BN. Обычно короткие связи приводят к ускорению сходимости и как минимум не снижают точность.

Визуальный анализ полученных результатов тестирования, в свою очередь, показал, что модификация U-Net с residual-блоками с задачей сегментации сетчатого ограждения справляется несколько лучше остальных моделей.

На рис. 4 можно увидеть, что модель U-Net с residual-блоками работает точнее и правильно распознает больше частей сетчатого ограждения, чем модель U-Net с энкодером VGG16BN, показавшая лучший результат по метрике IoU. При детальной визуальной оценке получаемых масок можно заметить, что U-Net с энкодером VGG16BN в сложных для распознавания местах не так четко прорисовывает нити сетчатого ограждения. На итоговых масках, полученных U-Net с residual-блоками, отсутствуют ложные порывы. Количество целых ячеек совпадает с числом целых ячеек на оригинальном

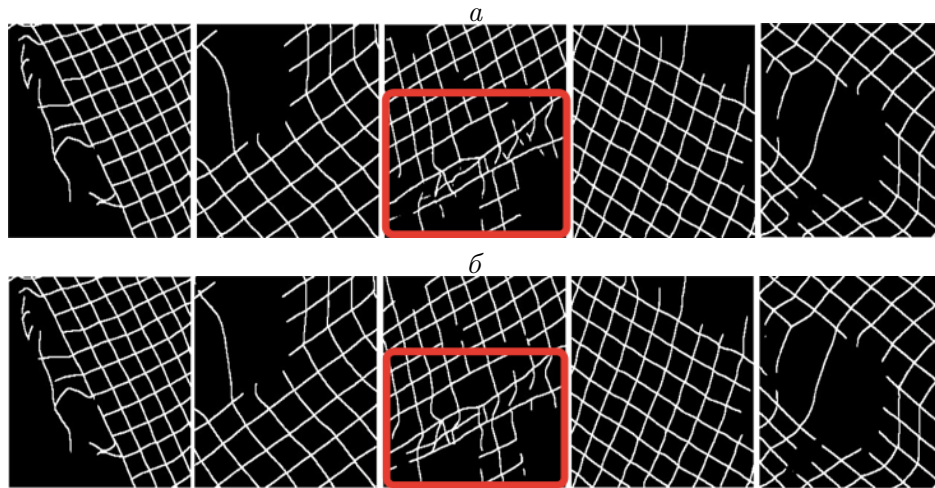


Рис. 4. Результаты семантической сегментации: *а* — модель с энкодером VGG16BN; *б* — модель с residual-блоками

Fig. 4. Results of semantic segmentation: *a* — a model with a VGG16BN encoder; *b* — a model with residual blocks

изображении. Отсутствие ложных порывов и ложных целых ячеек означает, что сгенерированные моделью на архитектуре U-Net с residual-блоками маски применимы для оценки состояния сетчатого ограждения.

3. Поиск повреждений сетчатого ограждения

Для решения задачи определения порывов в сеточных ограждениях разработаны и протестированы три алгоритма, основанные на классическом компьютерном зрении и использующие библиотеку OpenCV [22], а также модель машинного обучения на архитектуре U-Net. Для ее обучения создана специальная обучающая выборка, состоящая из размеченных масок изображений сетчатого ограждения с повреждениями. Все эти алгоритмы поиска повреждений [23] тестировались на выборке из 80 изображений. Факт присутствия повреждения и занятая им область определялись алгоритмами на основе классических методов компьютерного зрения с точностью от 88 до 95 % в зависимости от применяемого алгоритма. Модель машинного обучения на архитектуре U-Net продемонстрировала 100 %-ную точность, обнаружив все повреждения на тестовых изображениях. Тестирование показало, что модель машинного обучения на архитектуре U-Net определяет повреждения сетчатого ограждения точнее, чем алгоритмы на основе классического компьютерного зрения, но имеет более высокую вычислительную сложность.

На основе выполненных разработок алгоритмов компьютерного зрения и машинного обучения создан программный комплекс для поиска повреждений сетчатых ограждений подводных ферм. Его блок-схема представлена на рис. 5. Комплекс состоит из трех основных блоков, выполняющих предобработку изображений, их сегментацию и поиск повреждений. Блоки реализованы на языке программирования Python 3.7. Основные функции реализованы с использованием библиотек OpenCV 4.6.0, PyTorch 1.3.1, Skimage 0.13.8, Numpy 1.12.6, Matplotlib 3.2.2, TensorFlow 2.1.0, NVIDIA cuDNN 7.6.5.

Принцип организации работы блоков комплекса заключается в последовательной обработке ими цифровых изображений, полученных с фото- и видеокамер. Изображе-

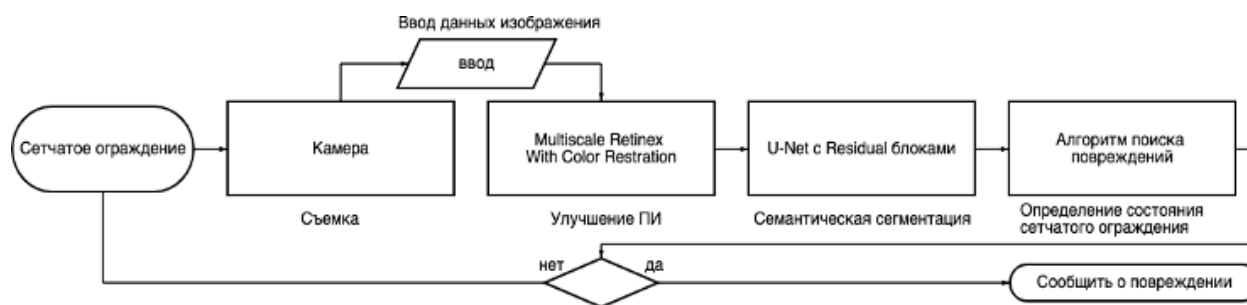


Рис. 5. Блок-схема программного комплекса

Fig. 5. Block diagram of the software package

ние с камеры попадает в блок улучшения изображения. В нем реализован алгоритм MSRCR и выполняется сжатие обработанного изображения до размера 256×256 пикселей. Затем оно передается в блок генерации бинарной маски, где реализована обученная модель глубокого обучения на модифицированной архитектуре U-Net с residual-блоками для семантической сегментации. Созданная после сегментации бинарная маска ограждения передается в третий (последний) блок, в котором происходит поиск повреждений ограждения.

При обнаружении повреждения на изображении соответствующая ему область выделяется и запускается алгоритм оповещения оператора. Если повреждений на маске не обнаружено, то исходное изображение снабжается соответствующей меткой, работа с ним завершается, а управление передается первому блоку. В первый блок комплекса подгружается следующее изображение, и запускается процесс его обработки.

В комплексе реализованы четыре упомянутых выше алгоритма поиска повреждений сетчатого ограждения. Выбор конкретного алгоритма может быть обусловлен реальными условиями эксплуатации АНПА и его вычислительными возможностями. Реализованные в этом комплексе алгоритмы поиска повреждений сетчатого ограждения, требования для их корректной работы и данные о точности распознавания ими повреждений, полученные при тестировании на выборке из 80 изображений, представлены в табл. 2.

Иллюстрация результата работы программного комплекса при определении повреждения сетчатого ограждения по фотоизображению приведена на рис. 6. В левой половине рисунка показаны исходное изображение и бинарная маска, полученная после его семантической сегментации. В правой представлены результаты наложения на исходное изображение его бинарной маски и области, содержащей обнаруженные поврежденные ячейки, которые окрашены в желтый цвет.

Т а б л и ц а 2. Требования и точность алгоритмов поиска повреждений

Table 2. Requirements and accuracy of damage detection algorithms

Метод	Камера	Реальные размеры	Точность, %
Сравнение периметров контуров	Стереокамера	Периметр целой ячейки	~89
Сравнение площадей ячеек	«	Площадь целой ячейки	~88
Комбинированный алгоритм	Монокамера	Не требуются	~95
Применение ИНС	«	«	100

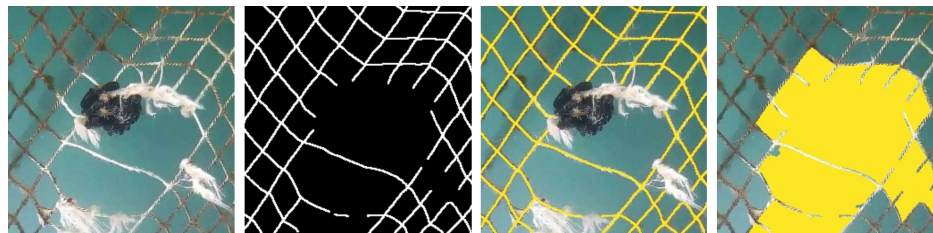


Рис. 6. Пример работы программного комплекса

Fig. 6. Example of the complex operation

Полная обработка комплексом одного изображения занимает ~ 0.22 с на персональном компьютере с процессором Intel Core i7-10700K, видеокартой MSI GeForce GTX 1070 Ti 8 и 16 ГБ оперативной памяти и SSD-диск. С учетом специфики задачи поиска повреждений и небольшой скорости движения АНПА нет необходимости в обработке всех 24 кадров видеопотока в секунду. Программный комплекс занимает 135 МБ на диске с учетом файлов весов моделей нейронной сети. Он может быть оптимизирован для работы с конкретным вычислительным оборудованием АНПА.

Заключение

Разработанный и реализованный в программном комплексе алгоритм предобработки изображений с помощью СНС корректно выделяет структурные объекты на подводных изображениях. Предложенная архитектура СНС и модель глубокого обучения на ее основе способны создавать фильтрующие маски сетчатого ограждения по снимкам, сделанным под водой. Разработанные алгоритмы компьютерного зрения и машинного обучения для оценки состояния сетчатого ограждения с использованием фильтрующей маски апробированы на реальных данных, полученных с борта ТНПА. Тестирование созданного программного комплекса показало возможность его использования в системах компьютерного зрения подводных роботов для определения состояния сетчатого ограждения в подводной среде по фотоизображениям.

Благодарности. Использованные в качестве исходных данных фотоизображения предоставлены коллегами из Китайского университета океанологии (г. Циндао). Для выполнения расчетов были использованы информационные и вычислительные ресурсы ЦКП “Центр данных ДВО РАН”.

Список литературы

- [1] Сергеев Л., Кузин В., Харин А., Мнацаканян Р., Мнацаканян А. Экономика рыбного хозяйства. Цифровизация управления: учебное пособие для вузов. М.: Издательство Юрайт; 2023: 318.
- [2] Colbourne D.B. Another perspective on challenges in open ocean aquaculture development. IEEE Journal of Oceanic Engineering. 2005: 30(1):4–11.
- [3] Huikai W., Junge Z., Kaiqi H., Kongming L., Yizhou Y. FastFCN: rethinking dilated convolution in the backbone for semantic segmentation. 2019: 1–15. Available at: <https://arxiv.org/pdf/1903.11816.pdf>.

- [4] **Garcia R., Nicosevici T., Cufi X.** On the way to solve lighting problems in underwater imaging. OCEANS'02 MTS/IEEE. IEEE; 2002; (2):1018–1024.
- [5] **Wang J., Lu K., Xue J., He N., Shao L.** Single image dehazing based on the physical model and MSRCR algorithm. IEEE Transactions on Circuits and Systems for Video Technology. 2017; 28(9):2190–2199.
- [6] **Смагин А.С., Дубровин К.Н.** Об алгоритмах компьютерного зрения для поиска порывов ветра в сетчатых ограждающих конструкциях. Вычислительные технологии. 2019; 24(6):118–125. DOI:10.25743/ICT.2019.24.6.014.
- [7] **Israfilov H.S.** Research of methods for binarization of images. Herald of Science and Education. 2017; 2(6):43–50.
- [8] **Guo Y., Liu Y., Georgiou T., Lew M.S.** A review of semantic segmentation using deep neural networks. International Journal of Multimedia Information Retrieval. 2018; (7):87–93.
- [9] **Zhou D., Fang J., Song X., Guan C., Yin J., Dai Y., Yang R.** Iou loss for 2d/3d object detection. 2019 International Conference on 3D Vision (3DV). 2019: 85–94.
- [10] **Качалин С.В.** Повышение стабильности обучения больших нейронных сетей путем дополнения небольших обучающих выборок родительских примеров синтезированными биометрическими примерами-потомками. Материалы научно-технической конференции кластера пензенских предприятий, обеспечивающих безопасность информационных технологий. Пенза; 2014; (9):32–35.
- [11] **Yaeger L., Lyon R., Webb B.** Effective Training of a neural network character classifier for word recognition. NIPS. 1996. Available at: <https://www.semanticscholar.org/paper/Effective-Training-of-a-Neural-Network-Character-Yaeger-Lyon/437fce6c281031a9dc69db9c54027b531dcbeec>.
- [12] **Ciresan D., Meier U., Gambardella L., Schmidhuber J.** Deep big simple neural nets excel on handwritten digit recognition. Neural Computation. 2010; 22(12). Available at: <https://arxiv.org/abs/1003.0358>.
- [13] **Simard P., Steinkraus D., Platt J.** Best practices for convolutional neural networks applied to visual document analysis. International Conference on Document Analysis and Recognition. IEEE; 2003. DOI:10.1109/ICDAR.2003.1227801. Available at: <https://ieeexplore.ieee.org/document/1227801>.
- [14] **Chakraverty S., Sahoo D.M., Mahato N.R.** McCulloch–Pitts neural network model. Concepts of Soft Computing. Springer; 2019: 167–173. Available at: https://link.springer.com/chapter/10.1007/978-981-13-7430-2_11.
- [15] **Ronneberger O., Fischer P., Brox T.** U-net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015: 18th International Conference. Springer International Publishing; 2015; 3(18):234–241.
- [16] **Badrinarayanan V., Kendall A., Cipolla R.** Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017; 39(12):2481–2495.
- [17] **He K., Gkioxari G., Dollar P., Girshick R.** Mask R-CNN. 2018; 1–12. Available at: <https://arxiv.org/pdf/1703.06870.pdf>.
- [18] **Wang Q., Ma Y., Zhao K., Tian Y.** A comprehensive survey of loss functions in machine learning. Annals of Data Science. 2022; (9):187–212. Available at: <https://link.springer.com/article/10.1007/s40745-020-00253-5>.

- [19] **Fathi A.** Semantic instance segmentation via deep metric learning. 2019. Available at: <https://arxiv.org/pdf/1703.10277.pdf>.
- [20] **Rosebrock A.** Intersection over Union (IoU) for object detection. 2016. Available at: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection>.
- [21] **Papandreou G.** Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. Proceedings of the IEEE International Conference on Computer Vision. 2015: 1742–1750.
- [22] **Bradski G.** The OpenCV library. Dr. Dobb's Journal: Software Tools for the Professional Programmer. 2000; 25(11):120–123.
- [23] **Смагин А.С.** Алгоритмы компьютерного зрения для оценки состояния подводных сетчатых ограждений. Дис. ... канд. техн. наук: 2.3.1. 2022: 145.

Automated search for damage to underwater mesh fences

SMAGIN SERGEY I.¹, SMAGIN ALEXEY S.^{2,*}

¹Computing Center of the FEB RAS, 680000, Khabarovsk, Russia

²Khabarovsk Federal Research Center of the FEB RAS, 680000, Khabarovsk, Russia

*Corresponding author: Smagin Alexey S., e-mail: smaginkhv@gmail.com

Received April 04, 2023, revised May 05, 2023, accepted May 12, 2023.

Abstract

Purpose. The aim of the work is to develop and study image processing algorithms designed to solve the problems of visual monitoring of the integrity and contamination of the mesh enclosing structures for underwater fish farms.

Methodology. The theoretical basis for the research relies on the methods and approaches used in the theory of pattern recognition, such as the analysis of the frequency characteristics of images using the Retinex transformation, the method of adaptive binarization of Otsu. Computer vision and machine learning technologies were used to develop algorithms for analyzing parametric contours and filtering masks of objects.

Findings.

1. New computer vision algorithms have been developed to assess the condition of mesh fences by their binary masks. This eliminates the processing of unnecessary data in the image, reduces time and increases the accuracy of damage detection.
2. A new modification of the convolution for neural network architecture has been developed for the problem of semantic segmentation of mesh fencing, the computational complexity of which is lower than that of the basic architectures described in the scientific literature. The proposed approach is based on the use of learnable regularization (residual blocks), which allows obtaining an object mask of higher quality than those obtained by classical computer vision.
3. A software package has been developed to automate the visual monitoring of underwater mesh fences, generating and using a binary fence mask to assess its condition.

Originality/value. The algorithms developed in the dissertation showed sufficiently high performance and significantly higher accuracy (from 88 to 100 %, depending on the algorithm used) for determining damage to underwater mesh fences than those proposed earlier in the works of other authors.

Keywords: automation, computer vision, machine learning, underwater vehicle, software package.

Citation: Smagin S.I., Smagin A.S. Automated search for damage to underwater mesh fences. Computational Technologies. 2023; 28(3):182–194. DOI:10.25743/ICT.2023.28.3.011. (In Russ.)

Acknowledgements. The photographic images used as initial data were provided by colleagues from the Chinese University of Oceanology (Qingdao). This research was supported in through computational resources provided by the Shared Services Center “Data Center of FEB RAS”

References

1. **Sergeev L., Kuzin V., Kharin A., Mnatsakanyan R., Mnatsakanyan A.** Ekonomika rybnogo khozyaystva. Tsifrovizatsiya upravleniya. Uchebnik dlya vuzov [Economics of fisheries. Digitalization of management. Textbook for universities]. Moscow: Izdatel'stvo Yurayt; 2023: 318. (In Russ.)
2. **Colbourne D.B.** Another perspective on challenges in open ocean aquaculture development. IEEE Journal of Oceanic Engineering. 2005: 30(1):4–11.
3. **Huikai W., Junge Z., Kaiqi H., Kongming L., Yizhou Y.** FastFCN: rethinking dilated convolution in the backbone for semantic segmentation. 2019: 1–15. Available at: <https://arxiv.org/pdf/1903.11816.pdf>.
4. **Garcia R., Nicosevici T., Cufi X.** On the way to solve lighting problems in underwater imaging. OCEANS'02 MTS/IEEE. IEEE; 2002; (2):1018–1024.
5. **Wang J., Lu K., Xue J., He N., Shao L.** Single image dehazing based on the physical model and MSRCR algorithm. IEEE Transactions on Circuits and Systems for Video Technology. 2017; 28(9):2190–2199.
6. **Smagin A.S., Dubrovin K.N.** On computer vision algorithms for searching breaks in meshed fencing constructions. Computational Technologies. 2019; 24(6):118–125. DOI:10.25743/ICT.2019.24.6.014. (In Russ.)
7. **Israfilov H.S.** Research of methods for binarization of images. Herald of Science and Education. 2017; 2(6):43–50.
8. **Guo Y., Liu Y., Georgiou T., Lew M.S.** A review of semantic segmentation using deep neural networks. International Journal of Multimedia Information Retrieval. 2018; (7):87–93.
9. **Zhou D., Fang J., Song X., Guan C., Yin J., Dai Y., Yang R.** Iou loss for 2d/3d object detection. 2019 International Conference on 3D Vision (3DV). 2019: 85–94.
10. **Kachalin S.V.** Increasing the learning stability of large neural networks by supplementing small training samples of parent examples with synthesized biometric descendant examples. Materialy Nauchno-Tekhnicheskoy Konferentsii Klastera Penzenskikh Predpriyatij, Obespechivayushchikh Bezopasnost' Informatsionnykh Tekhnologiy [Proceedings of the Scientific and Technical Conference of the Cluster of Penza Enterprises Providing Information Technology Security]. Penza; 2014; (9):32–35. (In Russ.)
11. **Yaeger L., Lyon R., Webb B.** Effective training of a neural network character classifier for word recognition. NIPS. 1996. Available at: <https://www.semanticscholar.org/paper/Effective-Training-of-a-Neural-Network-Character-Yaeger-Lyon/437fce6c281031a9dc69db9c54027b531dcbeecc>.
12. **Ciresan D., Meier U., Gambardella L., Schmidhuber J.** Deep big simple neural nets excel on handwritten digit recognition. Neural Computation. 2010; 22(12). Available at: <https://arxiv.org/abs/1003.0358>.
13. **Simard P., Steinkraus D., Platt J.** Best practices for convolutional neural networks applied to visual document analysis. International Conference on Document Analysis and Recognition. IEEE; 2003. DOI:10.1109/ICDAR.2003.1227801. Available at: <https://ieeexplore.ieee.org/document/1227801>.
14. **Chakraverty S., Sahoo D.M., Mahato N.R.** McCulloch–Pitts neural network model. Concepts of Soft Computing. Springer; 2019: 167–173. Available at: https://link.springer.com/chapter/10.1007/978-981-13-7430-2_11.

15. **Ronneberger O., Fischer P., Brox T.** U-net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015: 18th International Conference*. Springer International Publishing; 2015; 3(18):234–241.
16. **Badrinarayanan V., Kendall A., Cipolla R.** Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017; 39(12):2481–2495.
17. **He K., Gkioxari G., Dollar P., Girshick R.** Mask R-CNN. 2018; 1–12. Available at: <https://arxiv.org/pdf/1703.06870.pdf>.
18. **Wang Q., Ma Y., Zhao K., Tian Y.** A comprehensive survey of loss functions in machine learning. *Annals of Data Science*. 2022; (9):187–212. Available at: <https://link.springer.com/article/10.1007/s40745-020-00253-5>.
19. **Fathi A.** Semantic instance segmentation via deep metric learning. 2019. Available at: <https://arxiv.org/pdf/1703.10277.pdf>.
20. **Rosebrock A.** Intersection over Union (IoU) for object detection. 2016. Available at: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection>.
21. **Papandreou G.** Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. *Proceedings of the IEEE International Conference on Computer Vision*. 2015: 1742–1750.
22. **Bradski G.** The OpenCV library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*. 2000; 25(11):120–123.
23. **Smagin A.S.** Algoritmy komp'yuternogo zreniya dlya otsenki sostoyaniya podvodnykh setchatykh ograzhdeniy [Computer vision algorithms for assessing the condition of underwater mesh fences]. PhD Thesis: 2.3.1. 2022: 145. (In Russ.)