

О компьютерном зрении для мониторинга подводных ограждающих конструкций

А.С. Смагин

Хабаровский федеральный исследовательский центр ДВО РАН
Хабаровск, Россия
smaginkhv@gmail.com

Аннотация—Представлен метод автоматизации мониторинга состояния подводного сетчатого ограждения на примере предприятий аква-культурной направленности. Предложенный алгоритм выделения порывов сетчатых ограждений на основе искусственной нейронной сети показал хорошие результаты на реальных подводных снимках. Результаты, полученные программным комплексом, на основе созданного алгоритма приведены в статье.

Ключевые слова— алгоритм, компьютерное зрение, машинное обучение, программный комплекс.

1. ВВЕДЕНИЕ

Технологии компьютерного зрения и искусственные нейронные сети рассматриваются применительно к мониторингу состояния ограждающих конструкций подводных ферм на предприятиях аква-культурного профиля. Условия эксплуатации конструкций подводных ферм накладывают жесткие ограничения на непосредственное участие человека в их мониторинге и обслуживании. Поэтому разработка алгоритмов для систем компьютерного зрения подводных аппаратов, способных исключить или ограничить работу человека в сложных и опасных условиях подводного мира, является актуальной научно-практической задачей.

2. ПРЕДОБРАБОТКА ИЗОБРАЖЕНИЙ

Одной из основных задач компьютерного зрения является семантическая сегментация изображений [1], связанная с определением формы и положения объекта на изображении. Особую важность решение этой задачи имеет для определения характеристик аномальных областей рассматриваемых сетчатых ограждений, например, их дефектов [2].

На сложность выделения искомым объектов влияют такие факторы как: уровень яркости и его неравномерность, размытость изображения, низкая контрастность, а также посторонние объекты.

В разработанном программном комплексе поиск границ сетчатого ограждения выполняется двухэтапным алгоритмом предобработки изображений.

На первом этапе осуществляется компенсация артефактов, обусловленных освещением, с применением алгоритма Retinex [3, 4]. В его основе лежит идея разделения исходного изображения S на два, так чтобы яркость в каждой точке (x, y) рассчитывалась по формуле

$$S(x,y)=I(x,y)R(x,y).$$

Здесь I (illumination image) и R (reflectance image) - компоненты, отвечающие за освещение и отражение.

Используемая в программном комплексе модификация алгоритма Multi-Scale Retinex with Color Restoration хорошо приспосабливается к условиям освещения и справляется с задачей повышения яркости и контрастности изображения, проявляя скрытые из-за слабой освещённости и искажений детали изображения.

На втором этапе выполняется непосредственно семантическая сегментация. Используется модификация сверточной нейронной сети семейства R-CNN (Region Based Convolutional Neural Network), в частности Mask R-CNN [5].

В Mask R-CNN к традиционным для алгоритмов семейства R-CNN метке класса и координатам ограничивающей рамки добавляется также маска объекта — прямоугольная матрица принадлежности пикселя текущему объекту. Ключевым архитектурным отличием является применение RoIAlign вместо RoIPooling. RoIAlign не выполняет округлений сдвигов для пулинга, а сохраняет его значения с плавающей точкой, используя билинейную интерполяцию. Это обеспечивает более точное выделение маски объекта.

3. ОБУЧАЮЩАЯ ВЫБОРКА И ПРОЦЕСС ОБУЧЕНИЯ

Важной частью обучения нейронной сети является подготовка обучающей выборки. Главная проблема её подготовки состоит в сложности поиска нужных изображений, т. к. искомые объекты должны быть похожи на разных изображениях. Желательно, чтобы обучающая выборка состояла из большого числа изображений [6]. В разных источниках рекомендуемый размер обучающей выборки составляет примерно тысячу изображений и более. В нашем случае, в связи с отсутствием такого числа подходящих изображений, выборка состояла из 40 изображений. Разметка производилась с помощью инструмента VGG (Visual Geometry Group) Image Annotator. Из-за сложности структуры объекта разметка одного изображения занимала около часа. В дальнейшем планируется расширение данной выборки за счёт применения различных фильтров, искажающих картину.

В нашем случае, также, определяющее значение имеет, способ разметки. Каждое изображение необходимо поделить на более простые сегменты. Были протестированы 2 варианта разметки: с выделением сторон ячеек и их объединением в линии. Лучшие результаты распознавания были получены при использовании второго варианта разметки.

Обучение нейронной сети производилось на гибридном вычислительном кластере Вычислительного центра ДВО РАН. Использовались 2 10-ти ядерных

процессора IBM POWER8 (3.49 ГГц turbo, 160 потоков); память ECC, 256 ГБ; 2 NVIDIA Tesla P100 GPU.

Обучение сети происходило за 100 эпох, по 200 шагов в каждой и по 100 шагов оценки (validation steps). Длительность обучения ~5200 секунд. Распознавание объектов с помощью весов обученной нейронной сети зависит от размера изображения и не занимает много времени даже на относительно слабом оборудовании. Для изображений с разрешением 1920x1080 время распознавания составляло не более 5 секунд.

4. АЛГОРИТМ ПОИСКА ДЕФЕКТОВ ОГРАЖДАЮЩЕЙ КОНСТРУКЦИИ

После обработки изображения нейронной сетью мы получаем черно-белую маску изображения. Из-за низкого качества изображения или ошибок в работе нейронной сети на полученной маске могут появляться дефекты – порывы ячеек, которых на исходном изображении нет. На рисунке 1, слева внизу на изображении, показаны примеры таких порывов. Они выделены красными контурами.

Алгоритм работает следующим образом.

1. Выделение предполагаемой области дефекта. Предполагаемое место порыва определяется, сравнением числа пикселей изображений внутренних соседних ячеек (площадей ячеек). Пусть N общее число ячеек на изображении, N_i - множество номеров ячеек, соседних с ячейкой с номером i . Зададим числовые параметры $0 < a < b$ и обозначим через θ_{ij} отношение площадей ячеек с номерами i и j . Если $\forall ij, i \in N, j \in N_i$ выполняется условие $\theta_{ij} < a$, то переходим к п.5. Если $\exists ij, i \in N, j \in N_i$, что $\theta_{ij} > b$, то переходим к п.4. В остальных случаях переходим к п.2. Опытным путем установлено что, лучшие результаты достигаются при $a=1.5, b=2.5$.

2. Исключение ложных порывов. Для этого строятся две пары параллельных прямых, продолжающих стороны ячейки в месте порыва. Если угол между парами прямых меньше 20 градусов, то в этом месте может быть ложный порыв и переходим к п.3. Иначе - к п.4.

3. Восстановление маски сетчатого ограждения. Выполняется проверка: если длина предполагаемого дефекта занимает менее 0.2 длины стороны ячейки, то восстанавливаем маску в области дефекта путем закрашивания разрыва. Переходим к п.5. Если длина дефекта занимает более 0.2 от длины стороны ячейки, то делается вывод, что на маске изображен реальный порыв. Переходим к п.4.

4. Помечаем получившуюся область, как область реального порыва сетчатого ограждения. Переходим к анализу следующего изображению.

5. Отмечаем, что на рассмотренном изображении нет дефектов. Переходим к анализу следующему.

Пример области реального порыва представлен на рисунке 1 (верхнее правое изображение).

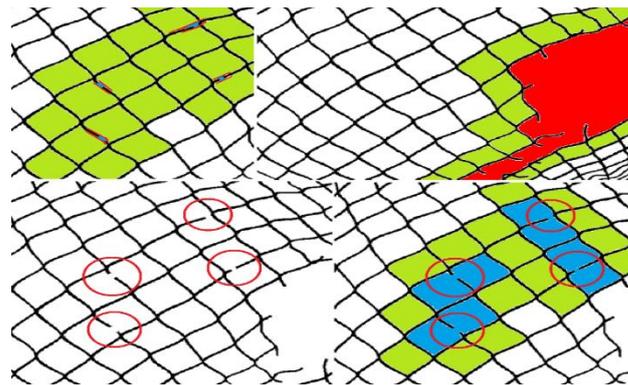


Рис. 1. Пример работы алгоритма

5. ЗАКЛЮЧЕНИЕ

Реализованный в программном комплексе подход к применению технологии компьютерного зрения при поиске порывов сетчатого ограждения, основанный на семантической сегментации, показал достаточно высокую эффективность и может стать основой для решения других сложных задач мониторинга ограждающих конструкций подводных ферм. Точность построения маски по метрике IoU (Intersection over Union) более 64%. Точность определения порывов 96%.

БЛАГОДАРНОСТИ

При выполнении исследования были использованы ресурсы ЦКП «Центр данных ДВО РАН».

ЛИТЕРАТУРА

- [1] Шапиро, Л. Компьютерное зрение (Computer Vision) / Л. Шапиро, Дж. Стокман. – М.: Бинوم. Лаборатория знаний, 2006. – 752 с.
- [2] Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, Ph. Fischer, Th. Brox // Medical Image Computing and Computer-Assisted Intervention. – 2015. – Vol. 9351. – P. 234-241.
- [3] Huikai, W. FastFCN: Rethinking Dilated Convolution in the Backbone for Semantic Segmentation / H. Wu, J. Zhang, K. Huang, K. Liang, Y. Yu. – 2019. – P. 1-15. – Access mode: <https://arxiv.org/pdf/1903.11816.pdf>.
- [4] Смагин, А.С. Об алгоритмах компьютерного зрения для поиска порывов в сетчатых ограждающих конструкциях / А.С. Смагин, К.Н. Дубровин // Вычислительные технологии. – 2019. – Т. 24, № 6. – С. 117-124.
- [5] Kaiming, H. Mask R-CNN / H. Kaiming, G. Gkioxari, P. Dollár, R. Girshick. – 2018. – P. 1-12. – Access mode: <https://arxiv.org/pdf/1703.06870.pdf>
- [6] Nordstrom, T. Using and designing massively parallel computers for artificial neural networks / T. Nordstrom, B. Svensson // Journal of Parallel and Distributed Computing. – 1992. – Vol. 14(3). – P. 260-285.