

Исследование структуры данных визуальной съёмки поверхности моря с применением методов машинного обучения

О. Белоусова

Московский физико-технический институт (МФТИ)
Институт океанологии им. П.П. Ширшова Российской
академии наук Москва, Россия
belousova.o@phystech.edu

М. А. Криницкий

Московский физико-технический институт (МФТИ)
Институт океанологии им. П.П. Ширшова Российской
академии наук Москва, Россия
krinitsky@sail.msk.ru

Аннотация—Обнаружение морского мусора является важной задачей для защиты экосистем Мирового океана, поскольку загрязнение угрожает биологическому разнообразию в гидросфере. Для решения данной проблемы в настоящем исследовании в качестве этапа разведочного анализа была изучена структура данных видеосъёмки поверхности моря при помощи искусственной нейронной сети, обучаемой в подходе контрастного обучения. Особенностью представленного нейросетевого алгоритма является применение специальной методики сэмплирования данных в процессе обучения. В будущем результаты данного исследования могут быть применены в задаче поиска аномальных объектов на поверхности моря.

Ключевые слова — *плавающий морской мусор, мониторинг состояния морской среды, искусственный интеллект, обнаружение объектов, машинное обучение, искусственные нейронные сети, исследование структуры в данных.*

I. ВВЕДЕНИЕ

Морской мусор – проблема, затрагивающая практически все регионы Мирового океана. Частицы мусора различных размеров, происхождения и состава сегодня встречаются практически во всех морях и на всех глубинах. Для решения проблемы загрязнения Мирового океана в первую очередь необходимо разработать и унифицировать методики наблюдений, способные дать объективные оценки плотности морского мусора и источников его происхождения. Один из наиболее доступных, при этом всё ещё объективных методов наблюдений - визуальная съёмка поверхности океана с последующим анализом полученных изображений. Цель нашего исследования – анализ изображений морской поверхности для последующего обнаружения экземпляров плавающего морского мусора. Среди интересующих нас явлений в данных могут быть: плавающий морской мусор, птицы, нестандартные блики. На настоящем этапе исследования мы проводим разведочный анализ данных фотоснимков морской поверхности, полученных с использованием фотокамеры высокого разрешения, закреплённой на леере морского исследовательского судна.

В основе нашей методики лежит метод обучения искусственной нейронной сети в подходе самоконтролируемого обучения, а именно контрастного обучения, что позволило проводить исследование без использования большого количества размеченных снимков. Объектами выборки в нашем исследовании считаются небольшие фрагменты снимков. Ключевой особенностью предлагаемого алгоритма обучения нейросети стал способ сэмплирования (извлечения элементов из выборки) фрагментов изображений поверхности моря, необходимых для контрастного самоконтролируемого обучения. В рамках нашего подхода сэмплирования мы предлагаем использовать свойство эргодичности поля морского

волнения и, как следствие, значительную пространственную автокорреляцию элементов изображения с существенным радиусом корреляции.

II. МЕТОДИКА ИССЛЕДОВАНИЯ И ИСХОДНЫЕ ДАННЫЕ

A. Данные

В качестве источника данных использовались видеозаписи, полученные в ходе научной экспедиции в Северном Ледовитом океане, которая проходила осенью 2023 года в Баренцевом и Карском морях. В ходе экспедиции велась видеосъёмка поверхности моря во время движения судна в светлое время суток. Камера устанавливалась на леере левого борта на высоте около 5,5 метров. Полоса видимости камеры составляла примерно от 6 до 15 метров в ширину и около 25 метров в длину от судна в зависимости от положения судна. Пример фотоснимка, полученного при разбиении видеозаписи на отдельные снимки приведён на рис.1.



Рис. 1. Пример фотоснимка с объектом плавающего морского мусора.

Для оценки количества аномалий, а также вычисления средних размеров аномалий и определения оптимальных размеров фрагментов была проведена визуальная разметка. Разметка набора данных проводилась при помощи приложения Label Studio. Искались следующие типы аномалий:

- птицы (Bird) – найдено 2716 объектов;
- морской мусор (Marine litter) – найдено 56 объектов;
- цветные блики (Glare) – найдено 1737 объектов.

Методика исследования

Цель исследования заключалась в поиске аномалий на снимках. Был выбран подход контрастного обучения (contrastive learning). Целью контрастного обучения является получение такого пространства скрытых представлений, в котором похожие пары образцов остаются близко друг к другу, а разнородные — далеко друг от друга. В настоящем исследовании размерность векторов скрытых представлений существенно ниже размерности признакового описания фрагментов.

В настоящем исследовании мы использовали нейросеть ResNet50, которую обучали методом MoCo (Momentum Contrast) [1] на объектах, представляющих собой небольшие фрагменты с характерными линейными размерами от 180x180 до 450x450 пикселей. В методе MoCo использовалась функция потерь InfoNCE:

$$\mathcal{L}_q = -\ln \frac{\exp\left(\frac{qk_+}{\tau}\right)}{\sum_{i=0}^K \exp\left(\frac{qk_i}{\tau}\right)}, \quad (1)$$

где q и k – векторы скрытого представления, вычисленные нейросетью на фрагментах изображения, причем, k_+ – вектор, соответствующий объекту, семантически совпадающему с объектом, соответствующим q согласно стратегии сэмплирования; остальные k_i считаются семантически различными с объектом, соответствующими q .

В подходе обучения MoCo для обучения модели необходимо создать массив так называемых положительных (q, k_+) и отрицательных (q, k_i) пар, скрытые представления которых участвуют в вычислении функции потерь. В нашем исследовании мы реализовали альтернативную по сравнению с оригинальной статьей MoCo [1] стратегию сэмплирования. В качестве положительной пары принимались два фрагмента из одной и той же фотографии, при этом расстояние между ними составляло не более 1 и не менее 0,2 длины опорного фрагмента. В качестве отрицательной пары использовались фрагменты с фотоснимков, отстоящие друг от друга на 10 или более кадров в рамках видеозаписи. При определении этих критериев мы руководствовались соображением, что соседние кадры отличаются друг от друга незначительно, а значит, семантически похожи друг на друга. Напротив, кадры из видеозаписей, полученных в разные дни, отличаются существенно.

После завершения обучения была применена процедура сокращения размерности с помощью UMAP [2] до размерности 2. Затем мы провели кластеризацию данных в форме двумерных векторов, используя метод Agglomerative Clustering [3].

III. РЕЗУЛЬТАТЫ

A. Функция потерь

Эволюция функции потерь отображена на рис. 2, на котором видно, что за 10 эпох обучения наблюдается устойчивый тренд на ее снижение. При этом следует отметить, что объём данных (количество фотографий, извлеченных из попутных видеозаписей) значителен и составляет 511262 полных снимков (см. рис. 1). Для повышения вычислительной эффективности и скорости стратегия сэмплирования положительных и отрицательных пар фрагментов снимка применялась таким образом, что с одного снимка были получены 200 пар фрагментов. Это приводит к тому, что объём данных, участвовавших в обучении на протяжении каждой эпохи, составляет 1022400 пар фрагментов.



Рис. 2. Эволюция функции потерь.

B. Распределение векторов скрытых представлений

На рис. 3 отображена двумерная диаграмма распределения векторов скрытых представлений из фрагментов изображений. График получен методом снижения размерности UMAP на 1/50 исходного набора данных, то есть примерно на 10000 полноразмерных фотоснимков.

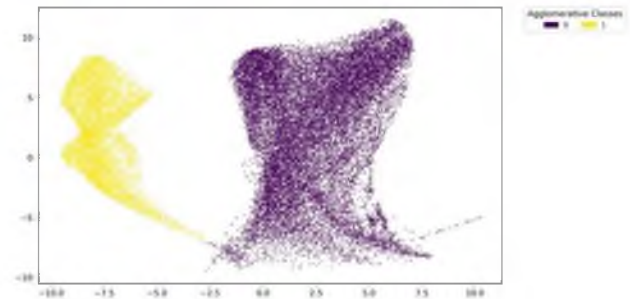


Рис. 3. Распределение векторов скрытых представлений.

Основной массив векторов (то есть, отдельных точек на диаграмме на рис. 3) однозначно выявляется моделью. В терминах результата кластеризации методом Agglomerative Clustering ему соответствует кластер «0». С другой стороны, модель кластеризации также выделяет и меньшие кластеры (номер «1» на рисунке), распознавая их отдельно от главного кластера. Метрика Mean Silhouette score (осредненный коэффициент силуэта) для приведенного результата кластеризации составила 0,47, что свидетельствует о высокой изоляции кластеров друг от друга. При этом на диаграмме рис. 3 видно, что имеются объекты, которые сложно однозначно отнести к какому-либо кластеру.

IV. ЗАКЛЮЧЕНИЕ

В рамках настоящего исследования мы провели анализ изображений морской поверхности при помощи подхода обучения нейросети MoCo, снижения размерности методом UMAP и кластеризации. Визуализация данных в форме распределения векторов сниженной размерности показывает наличие определённой структуры в данных, что можно в дальнейшем использовать в задаче идентификации аномалий в ходе решения задачи идентификации и классификации частиц плавающего морского мусора на поверхности моря без необходимости набора существенного объема размеченных снимков.

БЛАГОДАРНОСТИ

Настоящее исследование выполнено в рамках государственного задания по Соглашению № 075-03-2024- 117 от 17.01.2024.

ЛИТЕРАТУРА

- [1] He, K. Momentum Contrast for Unsupervised Visual Representation Learning / He, K., Fan, H., Wu, Yu., Xie, S., Girshick, R. // arXiv: <https://arxiv.org/abs/1911.05722>, 2020. – 1-12 p.
- [2] McInnes, L. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction / L. McInnes, J. Healy, J. Melville // <https://arxiv.org/pdf/1802.03426>, 2020 – 1-63 p.
- [3] Muellner, D. Modern hierarchical, agglomerative clustering algorithms / D. Muellner // <https://arxiv.org/pdf/1109.2378>, 2011 – 1-29 p.