

УДК 004.932.1:551.515.3

<http://doi.org/10.18799/29495407/2026/2/126>

Шифр специальности ВАК: 1.2.2

Научная статья



Влияние стратегий аннотации и масштаба входных данных на эффективность семантической сегментации океанических вихрей

И.А. Ботыгин^{1✉}, В.А. Крутиков², В.В. Кулак¹, В.С. Шерстнев¹, А.И. Шерстнева¹

¹Национальный исследовательский Томский политехнический университет,
Российская Федерация, г. Томск

²Институт мониторинга климатических и экологических систем СО РАН, Российская Федерация, г. Томск

✉bia@tpu.ru

Аннотация. Проведен аналитический обзор публикаций по семантической сегментации океанических вихрей. Показано, что архитектуры глубокого обучения, основанные на трансформерах, благодаря механизму самовнимания – наилучший вариант определения типа и общего контура вихря. Для оценки влияния стратегий аннотации и масштаба входных данных на качество сегментации с использованием модели SegFormer-B4 были сформированы наборы данных с масштабами 256×256, 512×512 и 1024×1024 пикселей. В экспериментах сравнивались два принципиально разных подхода к разметке изображений – детальная ручная аннотация (с филаментами) и упрощенная эллиптическая маска (без филаментов). Установлено, что модель, обученная на данных с включением спиральных филаментов, демонстрирует существенно более высокое качество сегментации по таким ключевым метрикам, как коэффициент Дайса (Dice Score) и индекс Жаккара (IoU). Выявлен оптимальный пространственный масштаб входных данных. Наилучшие показатели были достигнуты при использовании изображений размером 512×512 пикселей, что соответствовало исходному разрешению использованной предобученной модели SegFormer-B4. Сформулированы рекомендации по выбору стратегии аннотирования в зависимости от цели анализа – обнаружение наличия вихрей или точное восстановление их морфометрических характеристик.

Ключевые слова: семантическая сегментация, океанический вихрь, искусственные нейронные сети, машинное обучение, обработка изображений

Для цитирования: Ботыгин И.А., Крутиков В.А., Кулак В.В., Шерстнев В.С., Шерстнева А.И. Влияние стратегий аннотации и масштаба входных данных на эффективность семантической сегментации океанических вихрей. *Известия Томского политехнического университета. Промышленная кибернетика*, 2026, Т. 4, № 2, С. 35–47. <http://doi.org/10.18799/29495407/2026/2/126>

Вклад авторов: И.А. Ботыгин – формулирование идеи и цели исследования, написание первоначального текста статьи; В.А. Крутиков – научное руководство исследованием; В.В. Кулак – создание программного обеспечения, проведение экспериментов; В.С. Шерстнев – проверка результатов, проведение экспериментов, редактирование и доработка текста; А.И. Шерстнева – статистический анализ данных и результатов, разработка методов и процедур.

Использование средств ИИ: авторы заявляют о том, что при написании данной статьи не применялись средства генеративного искусственного интеллекта.

Конфликт интересов: отсутствует.

UDC 004.932.1:551.515.3

<http://doi.org/10.18799/29495407/2026/2/126>

Scientific paper

Impact of annotation strategies and input data scale on the efficiency of semantic segmentation of ocean eddies

I.A. Botygin^{1✉}, V.A. Krutikov², V.V. Kulak¹, V.S. Sherstnev¹, A.I. Sherstneva¹

¹ National Research Tomsk Polytechnic University, Tomsk, Russian Federation

² Institute of Monitoring of Climatic and Ecological Systems of the SB RAS, Tomsk, Russian Federation

✉bia@tpu.ru

Abstract. An analytical review of publications on the semantic segmentation of ocean eddies has been conducted. It was shown that transformer-based deep learning architectures, thanks to the self-attention mechanism, are the best option for determining the type and general contour of a eddies. To assess the impact of annotation strategies and input data scale on segmentation quality using the SegFormer-B4 model, the datasets with dimensions of 256×256, 512×512, and 1024×1024 pixels were created. The experiments compared two fundamentally different approaches to image annotation – detailed manual annotation (including filaments) and a simplified elliptical mask (excluding filaments). It was found that the model trained on data including spiral filaments demonstrated significantly higher segmentation quality according to key metrics such as the Dice score and the Intersection over Union (IoU). The authors identified the optimal spatial scale of the input data. The best results were achieved using images measuring 512×512 pixels, which corresponded to the native resolution of the pre-trained SegFormer-B4 model used. The authors formulated the recommendations regarding the choice of annotation strategy depending on the objective of the analysis – detecting the presence of vortices or accurately reconstructing their morphometric characteristics.

Keywords: semantic segmentation, ocean eddy, artificial neural networks, machine learning, image processing

For citation: Botygin I.A., Krutikov V.A., Kulak V.V., Sherstnev V.S., Sherstneva A.I. Impact of annotation strategies and input data scale on the efficiency of semantic segmentation of ocean eddies. *Bulletin of the Tomsk Polytechnic University. Industrial Cybernetics*, 2026, vol. 4, no. 2, pp. 35–47. <http://doi.org/10.18799/29495407/2026/2/126>

Authors' contribution: I.A. Botygin – statement of the idea and aim of the study, writing the original text of the article; V.A. Krutikov – scientific management of research; V.V. Kulak – creating software, conducting experiments; V.S. Sherstnev – verifying the results, conducting experiments, editing and refining the text; A.I. Sherstneva – statistical analysis of data and results, development of methods and procedures.

Use of AI: AI was not used.

Conflict of interest: none.

Введение

Семантическая сегментация изображений представляет собой одну из фундаментальных задач компьютерного зрения, целью которой является классификация каждого пикселя изображения в соответствии с его принадлежностью к определенному классу объектов. В отличие от классификации изображений, где определяется наличие объекта на всём изображении, или детекции объектов, где определяются их границы с помощью прямоугольников или эллипсов, семантическая сегментация обеспечивает точное «пиксельное» понимание объекта. То есть семантическая сегментация в данном случае спутниковых изображений океанических вихрей – это способ «понять» снимок на уровне каждого пикселя. Этот процесс позволяет не просто видеть океанические вихри, но и понимать их фор-

му, расположение и пространственные взаимоотношения, что является критически важным для оперативного, точного и воспроизводимого мониторинга их мезомасштабной динамики с целью физической параметризации.

Океанические мезомасштабные вихри (динамические образования) – распространённое и важное явление, влияющее на энергетический баланс, перенос тепла, солей, биогенных элементов, а также на формирование климата и крупномасштабной циркуляции вод. Их динамика особенно сложна для изучения в высоких широтах, где традиционные методы спутниковой альтиметрии малоэффективны из-за малого размера вихрей. Вихри играют важную роль в обмене теплом и газами между атмосферой и океаном, что критически важно для понимания глобальных климатических процессов.

Более точное их детектирование (обнаружение) и параметризация (сегментация) позволяют улучшить прогнозы погоды и подготовиться к стихийным бедствиям. Современные технологии детектирования океанических вихрей включают:

- *Спутниковые радиолокационные системы.* Используются спутниковые снимки, полученные с помощью радаров с синтезированной апертурой. Эти снимки позволяют выявлять вихри даже в сложных условиях, включая прикромочные ледовые зоны и открытую воду.
- *Искусственный интеллект и машинное обучение.* Нейросетевые алгоритмы используются для автоматического обнаружения и классификации вихрей на спутниковых снимках. Искусственный интеллект позволяет обрабатывать большие массивы данных, выявлять закономерности и строить прогнозы развития вихревых структур.
- *Численное моделирование и математические методы.* Разрабатываются математические модели, описывающие формирование и эволюцию вихрей в атмосфере и океане. Моделирование помогает понять внутреннюю структуру вихрей и их влияние на глобальные процессы.
- *Автономные датчики и буи.* Внедряются недорогие автономные датчики, которые перемещаются по океану и собирают данные о температуре, солёности, акустике и других параметрах. Эти устройства способны работать от 1 до 2 лет на стандартных батареях и до 10–20 лет в тяжёлых морских и метеорологических системах при регулярном обслуживании. Для последующего анализа данные регулярно передаются в облачные хранилища.
- *Интеграцию мультимодальных данных.* Объединяются данные с различных источников: спутниковые снимки, показания датчиков, акустические и визуальные данные. Это позволяет строить более точные и комплексные модели океанических процессов.

Развитие указанных технологий позволит получить более точные и оперативные сведения о динамике океана, что критически важно для науки и безопасности. Вместе с тем на практике возможны нюансы, которые будут влиять на качество аналитики:

- *Дисбаланс классов.* На изображении один класс может занимать гораздо больше площади, чем другой. Это сбивает нейросетевую модель при обучении.
- *Сложность границ.* Когда границы между объектами размыты, нейросетевой модели трудно точно их провести.
- *Разнообразие ракурсов и условий съёмки.* Спутники снимают под разными углами, в разное время суток и при разной освещённости – всё это влияет на внешний вид объектов.

- *Высокое разрешение.* Снимки высокого разрешения содержат много деталей, что может приводить к «шумовым» ошибкам.

Обзор публикаций по семантической сегментации океанических вихрей

Первые фундаментальные работы по сегментации изображений появились в середине 1960-х гг., когда зарождалась цифровая обработка сигналов и компьютерное зрение. В частности, основы базовых операций с пикселями при обработке цифровых изображений представлены в [1]. В начале 1970-х гг. появилась концепция растущих областей и пороговой фильтрации. Систематизация этих подходов изложена в [2]. А в конце 1970-х гг. был описан автоматический метод пороговой обработки изображений (глобальный метод бинаризации), который разделял пиксели на классы переднего и заднего плана путем максимизации дисперсии между ними [3]. Глобальное пороговое значение вычислялось непосредственно из гистограммы изображения. То есть оптимальный порог выбирался автоматически, максимизируя межклассовую дисперсию сегментированного изображения без ручной настройки. Переходом к семантической сегментации принято считать 2015 г., когда была опубликована архитектура полносверточной нейронной сети с эффективным выводом и обучением (Fully Convolutional Network, FCN) [4]. В FCN традиционные полносвязные слои были заменены сверточными слоями, что позволило нейросети принимать на вход изображения любого размера и выдавать плотные, попиксельные предсказания классов, соответствующие размерам исходного изображения.

В последние годы в связи с развитием глубокого машинного обучения семантическая сегментация активно применяется и для обработки спутниковых изображений, в частности океанических вихрей. В работе [5] для идентификации мезомасштабных антициклонических вихрей в море Лабрадор (Атлантический океан) представлен новый алгоритм, основанный на поиске локальных экстремумов относительной завихренности и сопутствующих геофизических полей. Для подбора оптимальных значений гиперпараметров в алгоритме идентификации вихрей используется байесовская оптимизация. В [6] исследуется архитектура глубокой сверточной нейронной сети (YOLOv5) на специально собранных и размеченных данных радиолокатора с синтезированной апертурой высокого разрешения для очень динамичной области над проливом Фрам (пролив между островом Гренландия и архипелагом Шпицберген). Сверточная нейронная сеть для обнаружения вихрей Ирмингера (образуются у западного побережья Гренландии из вод течения Ирмингера) представлена в [7]. Авторы исследуют два

подхода к обнаружению вихрей – традиционный алгоритм локальных экстремумов и модель глубокого обучения на основе U-Net. Для повышения точности обнаружения используется стратегия двухэтапного обучения, при которой нейронная сеть предварительно обучается на эвристических данных, а затем улучшается с помощью данных, размеченных экспертами. В [8] представлена гибридная архитектура нейросети, объединяющая классическую U-образную структуру кодировщика-декодера для сегментации изображений U-Net с остаточными блоками из сети ResNet. В качестве входных данных используются спутниковые данные о высоте уровня моря (Sea Surface Height, – SSH) и температуре поверхности моря (Sea Surface Temperature – SST). Те же данные SSH и SST с применением архитектуры U-Net используются и в [9–14]. В [15] в рамках семантической сегментации сравниваются результаты обнаружения вихрей с использованием традиционного динамического метода, сверточной нейросети PSPNet, нейросетевой архитектуры DeepLabV3+, архитектуры нейросети для семантической сегментации изображений в реальном времени BiSeNet. Обнаружено, что крупномасштабные вихри лучше выявляла архитектура BiSeNet, поскольку в нейросеть вводились направление, расстояние и геометрия для отображения движения вихря (пространственный путь). В [16] для идентификации мезомасштабных вихрей была использована та же нейросетевая архитектура DeepLabV3+. Для анализа были выбраны данные о высоте поверхности моря в Северной части Тихого океана со спутника Copernicus L4 за январь–июнь 2024 г. В [17, 18] для задачи идентификации и сегментации вихрей в хлорофилловых полях предложена многомасштабная архитектура нейронной сети MED-Net, в которой интегрируются данные цвета океана (параметр хлорофилл-а) с высоким разрешением, цифровая обработка изображений, искусственный интеллект и технологии многомасштабного обнаружения объектов. Облегченная модель CNN-Transformer под названием Ledformer для обнаружения вихрей описывается в [19]. Для повышения точности обнаружения используются данные об аномалиях уровня моря и данные о меридиональной и зональной скоростях вихря. В [20] описывается специализированная модель глубокого обучения Attention Dual-U-Net, которая использовалась для одновременного определения местоположения и контуров мезомасштабных вихрей в Южно-Китайском море. Модель интегрирует различные признаки, полученные из данных спутниковых наблюдений. В частности, абсолютную динамическую топографию и аномалии температуры поверхности моря для повышения точности обнаружения формы антициклонических и циклонических вихрей. Исследование с подоб-

ным набором данных, включающим абсолютную динамическую топографию, температуру поверхности моря и абсолютную геострофическую скорость, для изучения мезомасштабных вихрей северо-западной части Тихого океана приведено в [21]. В [22] предлагается подход к обнаружению вихрей, основанный на данных альтиметра, с использованием сети Ghost Attention Deep Lab Network (GAD-Net). Сеть Deep Lab с механизмом Ghost Attention сочетает в себе легковесные свертки Ghost, механизмы внимания и архитектуру семантической сегментации Deep Lab. Авторы декларируют, что GAD-Net обладает возможностями, сравнимыми с Py-Eddy-Tracker – инструментарием на Python для поиска, выделения и отслеживания мезомасштабных вихрей, превосходит другие модели по индексам оценки, имеет меньший размер модели и более низкую вычислительную сложность, а также обеспечивает обнаружение вихрей с меньшим радиусом и более высокой скоростью обнаружения. В [23] для идентификации океанических вихрей по данным дистанционного зондирования используется двухпирамидальная архитектура Dual-Pyramid UNet, которая сочетает модуль пирамидального разделенного внимания и пространственное пирамидальное пулинг-сложение с расширенным свертыванием. Кодировщик и декодер интегрируют низкоуровневые и высокоуровневые признаки, обеспечивая тем самым отсутствие значительной потери информации о признаках после нелинейного режима соединения. Результаты идентификации трехмерных структур мезомасштабных вихрей в Южном океане с использованием модели 3D-U-Res-Net представлены в [24, 25]. Обучение модели осуществлялось на наборе данных GLORYS12V1 за период с 2011 по 2020 гг. Вихри были классифицированы на четыре трехмерные структуры: чашеобразные, конусообразные, линзообразные и цилиндрические, с указанием их процентного соотношения и характеристик. В [26] предложена трехпараметрическая сверточная нейронная сеть для обнаружения мезомасштабных океанических вихрей. В нейросеть интегрированы три параметра: хлорофилл-а, коэффициент диффузного ослабления света в воде на длине волны 490 нм (сине-зеленый спектр) и температура поверхности моря. Каждый параметр проходит через отдельный путь кодирования-декодирования для получения карты признаков океанических вихрей на основе одного параметра. Затем три карты признаков объединяются и свертываются. После чего применяется заключительный слой активации для выполнения попиксельной классификации при обнаружении вихрей. Показано, что трехпараметрическая нейронная сеть демонстрирует значительно лучшие результаты, чем все однопараметрические модели. Достаточно полный обзор последних достижений (на 2024 г.) в

области идентификации океанических вихрей на основе искусственного интеллекта представлен в [27].

Анализ подходов к аннотации океанических вихрей

Исследование мезомасштабной изменчивости океана, центральное место в которой занимают океанические вихри, всё в большей степени опирается на синергию физического понимания процессов и методов машинного обучения. Несмотря на значительный прогресс в области машинной обработки, ключевым вызовом остается концептуальная проблема разметки (аннотирования) вихрей на спутниковых изображениях. Научная валидность и применимость выводов, полученных с помощью чувствительных и современных моделей нейросетей, напрямую зависят от принятой интерпретации границ вихревых структур.

Традиционно, начиная с основополагающих работ по созданию глобальных каталогов, вихри в полях SSH идентифицируются по замкнутым изолиниям, окружающим локальный экстремум (максимум для антициклона, минимум для циклона) [28]. На рис. 1 представлена сегментация объектов с выделением эллиптических масок.

Эллиптическая аннотация («без ножки») – вихрь выделяется как компактная замкнутая структура, приближенная к эллипсу. Такая разметка упрощает вычисления и повышает воспроизводимость, но может игнорировать физически значимую периферийную часть, так называемую «ножку», представляющую собой вытянутый градиент, или «хвост»

вихря, связанный с его динамикой и взаимодействием с окружающей средой. Эллиптическая аннотация базируется на строгой динамической интерпретации, где вихрь – это когерентная структура с компактным ядром, характеризующимся квази-геострофическим круговым движением.

Преимущества строгой аннотации носят фундаментальный характер для количественного анализа:

- Однозначность центра. Геометрический центр замкнутой контурной области (или центр, вычисленный через моменты распределения SSH) стабильно соответствует центру вращения.
- Физическая интерпретация метрик. Площадь, вычисленная внутри изолинии, прямо связана с кинетической энергией вихря, а его амплитуда легко определяется. Эти параметры согласуются с теоретическими моделями.
- Устойчивые траектории. Последовательность таких «строгих» центров формирует гладкую и физически правдоподобную траекторию, устойчивую к шуму и периферийным особенностям поля.

В пользу эллиптической аннотации можно указать, что такая разметка возможна как вручную, так и непосредственно с исходных спутниковых снимков. Сначала на изображении с помощью пороговой обработки или простого алгоритма кластеризации автоматически идентифицируются участки, которые потенциально могут быть вихрями. Затем к этим участкам применяются морфологические операции для удаления шума и устранения мелких артефактов.

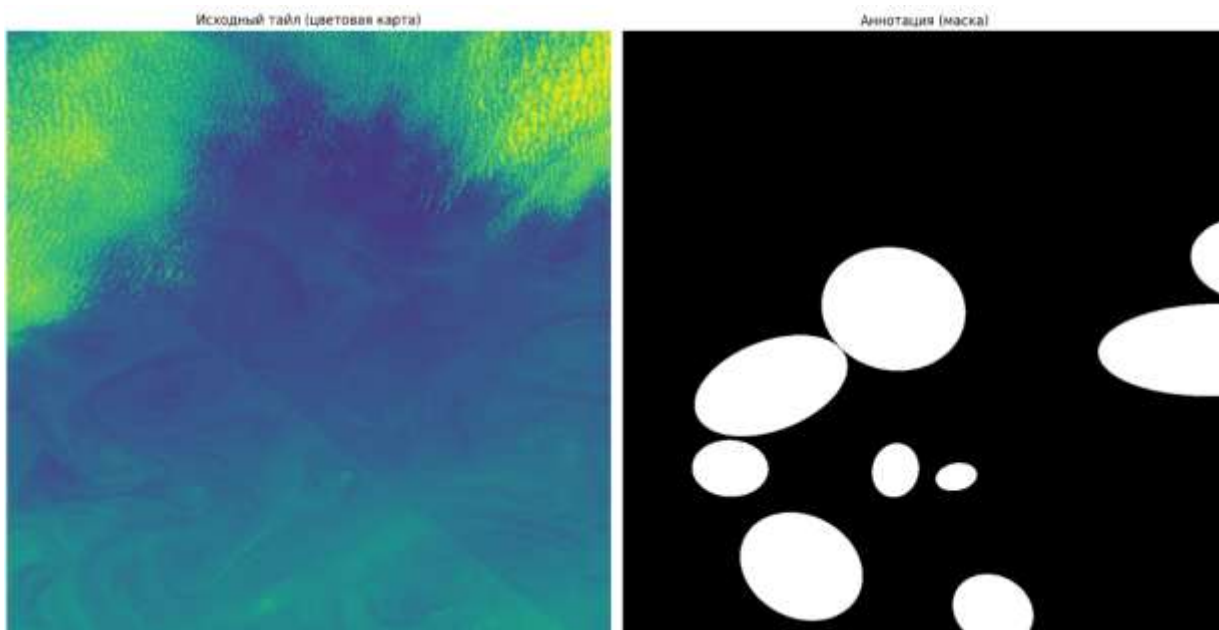


Рис. 1. Сегментация объектов: выделение эллиптических масок

Fig. 1. Object segmentation: elliptical mask extraction

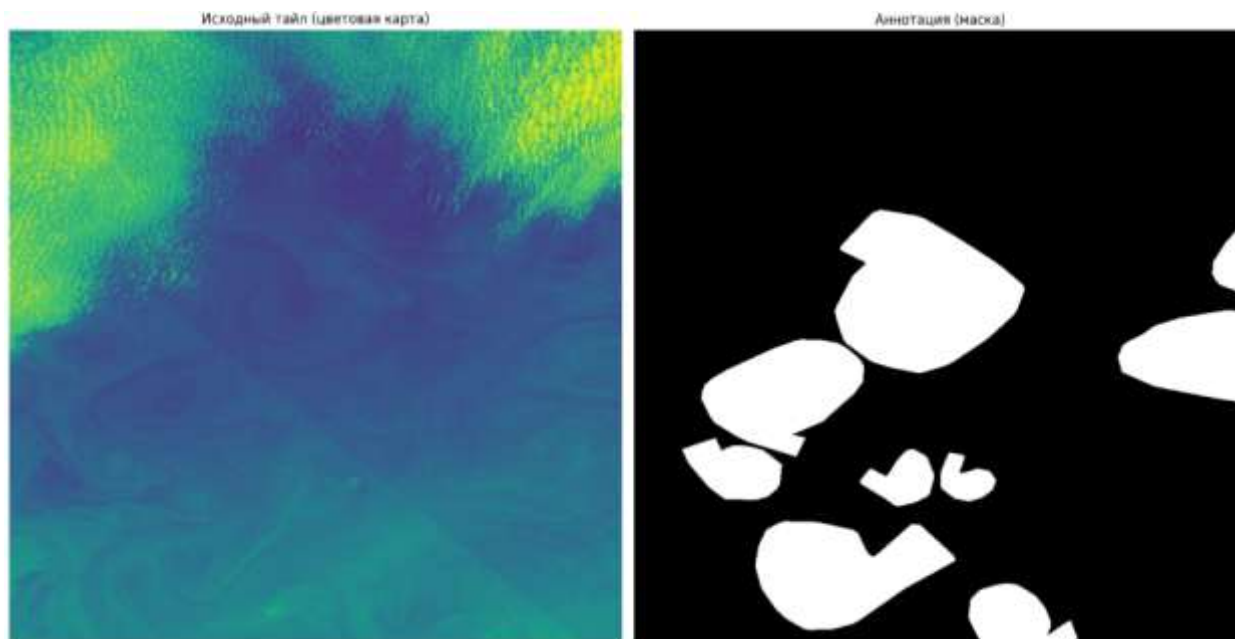


Рис. 2. Аннотация вихревых объектов: маскирование по контуру с включением ножки вихря
Fig. 2. Annotation of eddy objects: contour masking with the inclusion of an eddy «leg»

После этого выделяются крупные связные области и аппроксимируется каждая из них эллипсом, который наилучшим образом описывает форму выявленного объекта. Полученные эллипсы сохраняются как бинарные маски и могут использоваться для обучения модели. Этот способ особенно полезен, когда нет эксперта-океанолога для ручной разметки, но есть гипотеза о том, как вихри выглядят на спутниковых снимках.

Таким образом, данный подход является «консервативным», он гарантирует высокую достоверность обнаруженных объектов и корректность их физических характеристик, что критически важно для климатологии, расчета переноса свойств и верификации численных моделей.

Альтернативный подход фокусируется на морфологической целостности вихревого проявления [29]. Согласно этому подходу, вихрь часто наблюдается в данных не как идеальный замкнутый контур, а как ядро со спиральными филаментами («ножками»), которые являются динамическим следом его эволюции и взаимодействия со средой (рис. 2).

Полная контурная аннотация («с ножкой») – вихрь размечается вручную с учетом всех деталей морфологии, включая «ножку». Такой подход ближе к реальной структуре вихря, но менее унифицирован и более субъективен, поскольку проблема в том, что непонятно, где заканчивается эта структура. Например, разные эксперты могут по-разному интерпретировать слабо выраженные или частично маскированные участки. Один может включить в «ножку» тонкий градиентный хвост, другой – ограничиться только ярко выраженной частью.

Преимущества расширенной аннотации наиболее очевидны в контексте машинного обучения:

- Более богатый контекст для модели. Предоставление алгоритму не только ядра, но и характерных исходящих структур дает больше признаков для обучения, улучшая узнаваемость паттерна.
- Повышенная чувствительность. Модель, обученная на таких данных, может эффективнее обнаруживать слабовыраженные, формирующиеся или разрушающиеся вихри, а также структуры в сложных динамических областях.
- Отражение реальной морфологии. Такая разметка ближе к непосредственному визуальному восприятию вихревых структур на растрах высоты поверхности моря или картах скоростей.

Таким образом, в отличие от эллиптической аннотации, где есть четкий математический алгоритм (например, аппроксимация по центру масс и главным осям), для «ножки» нет универсального правила. Это делает процесс разметки не автоматизируемым и трудно воспроизводимым. Также можно подчеркнуть зависимость от качества данных, когда на изображениях с атмосферными помехами (облачность) или шумом граница «ножки» может быть практически не различима. В таких случаях эксперт-океанолог вынужден принимать решение на основе интуиции, что еще больше увеличивает вариативность. Неоднородность морфологии выражается в том, что океанические вихри имеют очень разную форму – одни круглые, другие сильно вытянутые, некоторые сливаются друг с другом. При разметке «с ножкой» для каждого вихря тре-

буется индивидуальный подход, что затрудняет стандартизацию процедуры.

Ключевой недостаток этого подхода – компромисс с физической метрикой. Включение филаментов приводит к систематическому завышению площади, смещению вычисленного центра тяжести и, как следствие, к появлению шума в определяемых траекториях. Это делает прямую количественную интерпретацию выходных данных модели проблематичной. Возникает противоречие между эффективностью обнаружения (которой способствует расширенная аннотация) и точностью количественных оценок (которая обеспечивается строгой аннотацией). С одной стороны, для точного физического анализа (расчета энергии, площади, траектории) необходима разметка, отражающая только ядро вихря, а с другой, для эффективного автоматического обнаружения, предпочтительна «расширенная» разметка, включающая филаменты. Несмотря на большую физическую достоверность, аннотация «с ножкой» менее пригодна для обучения моделей нейросетей из-за высокой субъективности и отсутствия единых критериев разметки. Выбор стратегии разметки становится не техническим, а методологическим решением, определяющим применимость модели для интерпретации результатов искусственных нейросетей в области физической океанографии.

Модель SegFormer-B4

До недавнего времени в задачах семантической сегментации лидировали сверточные нейронные сети (Convolutional Neural Network – CNN) и их производные. Но в последние годы архитектуры глубокого обучения, основанные на трансформерах, постепенно начинают доминировать. В частности, модель SegFormer сочетает в себе эффективность сверточных нейронных сетей и способность трансформеров улавливать глобальные зависимости и благодаря своей иерархической структуре и эффективному механизму агрегирования контекста показывает высокую точность при одновременной вычислительной экономичности, что делает её перспективным инструментом для анализа океанологических изображений [30, 31].

Модель SegFormer имеет четыре основные модификации (B0÷B5), где буква «B» означает Base, а цифра указывает на масштаб – от самого легкого варианта (B0) до самой сложной версии (B5). Вариант SegFormer-B4 – баланс между точностью и скоростью (~64,1 млн параметров). Модель заметно компактнее и быстрее многих CNN-решений с сопоставимой точностью.

Модель SegFormer B4 состоит из двух основных частей:

- Энкодер (кодировщик) Mix Transformer (MiT-B4). В отличие от классических Vision

Transformer, которые работают с изображениями фиксированного разрешения, MiT-B4 сразу генерирует иерархические признаки на нескольких масштабах (H/4, H/8, H/16, H/32) – постепенно уменьшает пространственное разрешение и увеличивает размерность каналов, как это делают CNN. Это важно для сегментации – нужны и крупные контексты, и мелкие детали.

- Декодер на основе полносвязного многослойного персептрона. Имеет только линейные слои. Декодер берёт признаки с разных уровней энкодера, выравнивает каналы, масштабирует до нужного разрешения, объединяет и выдаёт финальную карту сегментации.

Важные особенности модели SegFormer B4:

- Глобальный контекст. Трансформер естественным образом моделирует длинные зависимости. Это помогает понимать структуру крупных вихревых систем.
- Иерархия признаков. MiT-B4 даёт признаки на разных уровнях детализации. Можно одновременно сегментировать и крупные вихри, и мелкие спиральные структуры.
- Перекрывающиеся фрагменты. Чтобы не терять локальную непрерывность на границах фрагментов, используется вложение с перекрытием. Это помогает лучше сохранять границы вихрей.
- Гибкость по разрешению. Отсутствие позиционных кодировок упрощает работу с тайлами разного размера и с разными масштабами. Удобно при подготовке набора данных из большого количества тайлов с разными размерами и кадрированием, а также когда на обучении и инференсе изображения разного размера.
- Многоканальность входа. Совместимость с данными радаров с синтезированной апертурой при дообучении.

Практическая реализация сегментации океанических вихрей

В задачах сегментации океанических вихрей по спутниковым данным эффективность использования нейросетевых моделей критически зависит от качества обучающих аннотаций. Выше было отмечено, что в этих задачах долгое время доминировал физически обоснованный подход, при котором границы вихря проводятся по замкнутым изолиниям, ограничивающим только компактное вращающееся ядро, а вытянутые спиральные филаменты («ножки») рассматриваются как артефакт и отсекаются. Современные представления о динамике вихрей рассматривают спиральные филаменты как неотъемлемую часть протяженного составного объекта, переносящего воду и вещества из периферийной зоны к центру. Учет этих структур расширяет пространственный контекст, улучшает устойчивость алгоритмов к зашумленности спутниковых

данных и повышает точность восстановления контура вихря в целом.

В настоящей работе в контексте бинарной сегментации океанических вихрей проведено экспериментальное сравнение двух стратегий аннотирования – с филаментами (маска «ножки» – с включением артефактов вихря) и без филаментов (маска «эллипс» – упрощенная эллиптическая форма). С использованием модели SegFormer-B4 проведены исследования по поиску баланса между точностью обнаружения и физической корректностью результирующих масок. В эксперименте одна и та же архитектура модели SegFormer-B4, предварительно обученная на сценарном наборе данных ADE20K, последовательно дообучалась на трех вариантах входных тайлов: 256×256, 512×512 и 1024×1024 пикселей. Такой подход позволил не только оценить, как тип аннотации влияет на качество сегментации, но и выявить, насколько модель чувствительна к масштабу представления данных при разных стратегиях разметки.

Для экспериментов был подготовлен набор спутниковых изображений, охватывающий Лофотенскую котловину (Норвежское море) за 2017–2025 гг.

Первый эксперимент: три группы данных с концентрацией хлорофилла-а, сформированные под маски с филаментами:

- 197 тайлов под обучение и 68 для теста, формат GeoTIFF, 1024×1024 пикселей;
- 755 тайлов под обучение и 238 для теста, формат GeoTIFF, 512×512 пикселей;
- 562 тайла под обучение и 172 для теста, формат GeoTIFF, 256×256 пикселей.

Разметка масок с филаментами осуществлялась вручную по контуру с включением периферийных «хвостов» вихрей на основе экспертной визуальной интерпретации.

Второй эксперимент: три группы данных с концентрацией хлорофилла-а, сформированные под маски «эллипс»:

- 170 тайлов под обучение и 65 для теста, формат GeoTIFF, 1024×1024 пикселей;
- 683 тайла под обучение и 225 для теста, формат GeoTIFF, 512×512 пикселей;
- 527 тайлов под обучение и 178 для теста, формат GeoTIFF, 256×256 пикселей.

Маски «эллипс» – это те же изображения, но аннотированные путем аппроксимации вихря эллипсом (по центру масс и главным осям структуры). Все детали, не входящие в эллипс (в т. ч. филаменты), были отброшены.

Все наборы данных были разделены на обучающую и валидационную выборки в соотношении 80/20. Для обучения использовались стандартные расширения наборов данных (аугментации): случайное отражение (горизонтальное/вертикальное),

нормализация изображений по параметрам набора ImageNet-1K (среднее значение $\mu=[0,485, 0,456, 0,406]$ и стандартное отклонение $\sigma=[0,229, 0,224, 0,225]$ для каналов R, G, B соответственно), масштабирование. При изменении размера тайлов (256×256, 512×512, 1024×1024 пикселей) изображения и маски масштабировались с сохранением соотношения сторон, без искажений.

Для объективной оценки качества автоматического выделения вихрей использовались следующие стандартные метрики:

- *Показатель Dice Score (F1).* Dice Score (коэффициент Дайса) и F1-score (F1-мера/гармоническое среднее) – математически эквивалентные статистические показатели. Используются для измерения совпадения/сходства между наборами данных. Коэффициент Дайса измеряет пространственное перекрытие между прогнозируемой маской изображения и эталонным изображением (пересечение, деленное на полусумму размеров). Значение варьируется от 0 (отсутствие перекрытия) до 1 (идеальное совпадение). F1-мера – это показатель прогностической эффективности. Рассчитывается как гармоническое среднее точности (Precision) и полноты (Recall), симметрично отражая их в одной метрике. Максимально возможное значение F-меры 1, что указывает на идеальную точность и полноту, а минимально возможное значение 0, если точность или полнота равны нулю. Для задач бинарной классификации они эквивалентны и на уровне пикселей дают абсолютно одинаковое числовое значение. Обозначаем обобщенную метрику как «Показатель Dice Score (F1)».
- *Показатель IoU (индекс Жаккара).* Измеряет, насколько хорошо предсказанная моделью форма или граница совпадает с фактическим целевым объектом – делит размер пересечения на размер объединения множеств.
- *Accuracy (доля правильных ответов).* Отношение верных предсказаний ко всем примерам.
- *Precision (точность).* Сколько предсказанных положительных элементов оказались правильными – положительная прогностическая ценность.
- *Recall (полнота/чувствительность).* Доля найденных положительных объектов от их реального общего числа в данных.

При расчете метрик качества были учтены крайние случаи, например, ситуации, когда маска или предсказание модели оказываются полностью пустыми (отсутствие вихрей на изображении). Соответствующая обработка таких случаев предотвращала вычислительные ошибки и обеспечивала корректность количественной оценки.

Эксперименты проводились на трех масштабах тайлов – 256×256, 512×512 и 1024×1024 пикселей.

Для каждого масштаба оценивалось два типа разметки – с учетом филаментов вихрей и без них. Значения полученных в результате экспериментов метрик качества выделения океанических вихрей с использованием модели SegFormer-B4 представлены в таблице.

Таблица. Метрики качества автоматического выделения вихрей

Table. Quality metrics for automatic eddy detection

Размер изображений (пиксели) Image size (pixels)	Доля правильных ответов (Accuracy) Percentage of correct answers (Accuracy)	Точность (Precision)	Полнота/ чувствительность (Recall) Fullness/Sensitivity (Recall)	Показатель Dice Score (F1)	Индекс Жаккара (IoU) Jaccard index (IoU)
С филаментами/With filaments					
1024×1024	0,88	0,63	0,29	0,40	0,25
512×512	0,98	0,94	0,91	0,93	0,87
256×256	0,90	0,81	0,81	0,81	0,68
Без филаментов/Without filaments					
1024×1024	0,88	0,62	0,23	0,35	0,22
512×512	0,93	0,82	0,81	0,82	0,69
256×256	0,77	0,45	0,21	0,29	0,17

Для упрощения визуального сравнения результатов бинарные маски отображались в черно-белом представлении: фон кодировался черным цветом, а область вихря – белым. Такой подход минимизировал когнитивную нагрузку на пользователя и позволял оценить соответствие предсказания эталонной разметке.

Обсуждение полученных результатов

В ходе проведения серии экспериментов было установлено, что выбор формы аннотации оказывает существенное влияние на качество обучения модели семантической сегментации океанических вихрей. В частности, модели, обученные на разметке с филаментами, демонстрируют стабильно лучшие результаты по сравнению с теми, что обучались на упрощенных эллиптических масках. Это различие обусловлено несколькими взаимосвязанными факторами. Прежде всего, аннотации с филаментами обладают большей физической адекватностью. Реальные вихри редко имеют идеально симметричную эллиптическую форму, их структура часто включает асимметричные, вытянутые или размытые периферийные участки, возникающие в результате взаимодействия с окружающей средой и внутренней динамикой потока. Эллиптическая маска, напротив, представляет собой геометрически упрощенную аппроксимацию, которая не отражает истинной сложности объекта и может

включать в разметку области, не относящиеся к вихрю, либо игнорировать его физически значимые границы. Кроме того, разметка с филаментами заставит модель учиться более точному воспроизведению сложных нерегулярных границ, включая тонкие структуры и плавные переходы между вихрем и фоном.

Модель, обученная на эллипсах, склонна к формированию обобщенных гладких предсказаний, которые могут демонстрировать высокие значения индекса Жаккара (IoU) за счет «заливки» центральной области, но при этом терпят неудачу в периферийных зонах, критичных для научной интерпретации.

Важную роль также играет семантическая насыщенность разметки. Аннотация с филаментами – это не просто двоичная маска, а результат осознанного экспертного решения, отражающего динамическую природу явления. Такая разметка передает не только пространственное положение объекта, но и его структурные особенности, что повышает обобщающую способность модели на данных, содержащих реалистичные (неидеальные) вихри.

Таким образом, несмотря на большую трудоемкость, аннотации с филаментами обеспечивают более глубокое и физически обоснованное обучение, что приводит к улучшению качества сегментации. Полученные результаты подчеркивают, что в задачах, где объекты обладают сложной морфологией и нестационарной структурой, качество и семантическая точность разметки важнее её формальной простоты. Следовательно, при формировании наборов данных для задач сегментации всех геофизических явлений (не только океанических вихрей) предпочтение следует отдавать аннотациям, максимально приближенным к физической реальности, даже если они требуют значительных усилий при ручной разметке.

Поскольку модель SegFormer-B4 изначально обучалась на наборе данных ADE20K при входном разрешении 512×512 пикселей, этот размер является априори оптимальным с точки зрения архитектурных предпосылок и распределения признаков в сверточных слоях. Поэтому в ходе обучения все входные тайлы, независимо от исходного размера, были масштабированы до 512×512 пикселей, чтобы соответствовать входному формату модели. Но, как показали результаты экспериментов, качество сегментации существенно зависело от исходного размера тайла (размера до масштабирования), то есть от того, насколько детализированным было исходное изображение.

Наилучшее качество сегментации достигается при использовании тайлов размером 512×512 пикселей, что полностью соответствует исходному разрешению, на котором была обучена модель. На этом масштабе модель демонстрирует максималь-

ные значения всех ключевых метрик: Precision=0,94, Recall=0,91, Dice Score (F1)=0,93, IoU=0,87 для разметки с филаментами. Это указывает на способность модели точно выделять не только основную область вихря, но и его мелкие структурные элементы, так называемые филаменты, которые требуют достаточного пространственного разрешения для корректного распознавания.

При уменьшении исходного размера тайла до 256×256 пикселей наблюдается снижение всех метрик: Dice Score (F1) падает до 0,81, а IoU – до 0,68. Это связано с тем, что при масштабировании с 256×256 до 512×512 пикселей происходит искусственное увеличение изображения, которое не восстанавливает потерянные детали, а лишь интерполирует их, что снижает точность и полноту предсказаний.

При увеличении исходного размера тайла до 1024×1024 пикселей, несмотря на более высокое разрешение, качество сегментации не улучшается, а даже снижается, что особенно заметно в случае разметки без филаментов, где IoU падает до 0,22, а Dice Score (F1) – до 0,35. Этот эффект обусловлен тем, что при масштабировании с 1024×1024 до 512×512 пикселей происходит потеря локальной информации, что нарушает пространственные соотношения между объектами

Теоретически показатель Dice Score (F1) монотонно связан с индексом Жаккара (IoU), но при аннотировании с филаментами Dice Score (F1) даёт более высокое числовое значение, чем индекс Жаккара (IoU), для одного и того же прогноза. Это объясняется тем, что Dice Score (F1) придает больший вес зоне перекрытия. При аннотировании упрощенной эллиптической формой числовые значения Dice Score (F1) и индекса Жаккара (IoU) ближе друг к другу, поскольку маски компактны и четко определены.

Таким образом, результаты экспериментов подтверждают, что выбор стратегии аннотации должен соответствовать цели исследования:

- Если цель – максимально точно описать всю вихревую структуру (например, для оценки площади или энергии), то предпочтительна раз-

метка с филаментами, несмотря на большую вариативность границ.

- Если цель – автоматизировать регистрацию ядра вихря (например, для определения центра или ориентации), то эллиптическая аннотация обеспечивает более стабильные и воспроизводимые результаты.

Заключение

В контексте бинарной сегментации океанических вихрей с использованием модели SegFormer-B4 проведены эксперименты, в которых сравнивались два принципиально разных подхода к разметке – детальное ручное аннотирование «с ножкой» и упрощенная эллиптическая маска («без ножки») на трех разрешениях входных тайлов – 256×256 , 512×512 и 1024×1024 пикселей.

Сделан вывод, что выбор стратегии разметки должен быть тесно увязан с конечной задачей. Расширенное аннотирование предпочтительно для задач детектирования и качественного анализа структуры, тогда как строгая эллиптическая – для точного измерения параметров ядра вихря.

Для комплексного решения задач обнаружения и количественного анализа предлагается комбинированная стратегия – обучение модели сегментации проводить на данных с расширенным аннотированием («с ножкой») для максимизации эффективности обнаружения и устойчивости к шуму, а на этапе постобработки результатов применять алгоритмы выделения компактного ядра (эллиптическую аппроксимацию) для получения физически корректных метрик. Такой подход позволяет совместить преимущества обеих стратегий аннотирования.

Полученные выводы и гибридная методика, где эллипс используется для первоначальной локализации центра, а полная сегментация для уточнения границ, могут быть применены не только для сегментации океанических вихрей, но и для анализа других классов динамических природных объектов, где существует схожий компромисс между морфологической целостностью и метрической точностью.

СПИСОК ЛИТЕРАТУРЫ

1. Rosenfeld A., Pfaltz J.L. Sequential operations in digital picture processing. *Journal of the ACM*, 1966, Vol. 13, Iss. 4, P. 471–494. DOI: <https://doi.org/10.1145/321356.321357>
2. Zucker S.W. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 1976, Vol. 5, Iss. 3, P. 382–399. DOI: [https://doi.org/10.1016/S0146-664X\(76\)80014-7](https://doi.org/10.1016/S0146-664X(76)80014-7)
3. Otsu N. A Threshold selection method from gray-level histograms. *IEEE Transaction on Systems, Man and Cybernetics*, 1979, Vol. 9, Iss. 1, P. 62–66. DOI: 10.1109/TSMC.1979.4310076
4. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA, 2015. P. 3431–3440. DOI: 10.1109/CVPR.2015.7298965
5. Калинин М.А., Вереземская П.С., Криницкий М.А., Борисов М.А., Тилинина Н.Д. Байесовская оптимизация метода идентификации мезомасштабных океанских вихрей в море Лабрадор в данных вихреразрешающего моделирования. *Океанологические исследования*, 2024, Т. 52, № 4, С. 56–73. DOI: 10.29006/1564-2291.JOR-2024.52(4).4 EDN: FSRROH
6. Khachatrian E., Sandalyuk N., Loizou P. Eddy detection in the marginal ice zone with Sentinel-1 data using YOLOv5. *Remote Sensing*, 2023, Vol. 15, Iss. 9, art. 2244. DOI: <https://doi.org/10.3390/rs15092244> EDN: GGCCEI

7. Kalinin M.A., Krinitskiy M.A., Verezemskaya P.S. Detection of Irminger rings in high resolution ocean hydrodynamic modeling data using artificial neural networks. *Moscow University Physics Bulletin*, 2025, Vol. 80, Suppl 3, P. S1270–S1278. DOI: <https://doi.org/10.3103/S0027134925703102> EDN: PMSWYV
8. Liu Y., Zheng Q., Li X. Detection and analysis of mesoscale eddies based on deep learning. *Artificial intelligence oceanography*. Eds. Li X., F. Wang. Singapore: Springer, 2023. P. 209–225. DOI: https://doi.org/10.1007/978-981-19-6375-9_10
9. Liu Y., Li X., Ren Y. A deep learning model for oceanic mesoscale eddy detection based on multi-source remote sensing imagery. *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, P. 6762–6765. DOI: 10.1109/IGARSS39084.2020.9323716
10. Geng J., Gao H., Huang B., Radenkovic M., Chen G. ARU²-net: a deep learning approach for global-scale oceanic eddy detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, Vol. 17, P. 11997–12007. DOI: 10.1109/jstars.2024.3419175 EDN: AQBZOB
11. Sun X., Zhang M., Dong J., Lguensat R., Yang Y., Lu X. A deep framework for eddy detection and tracking from satellite sea surface height data. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, Vol. 59, № 9, P. 7224–7234. DOI: 10.1109/TGRS.2020.3032523 EDN: BZLKWB
12. Zhao Y., Fan Z., Li H., Zhang R., Xiang W., Wang S., Zhong G. SymmetricNet: end-to-end mesoscale eddy detection with multi-modal data fusion. *Frontiers in Marine Science*, 2023, Vol. 10, art. 1174818. DOI: <https://doi.org/10.3389/fmars.2023.1174818> EDN: ERMINB
13. Huo J., Zhang J., Yang Ju., Li Ch., Liu G., Cui W. High kinetic energy mesoscale eddy identification based on multi-task learning and multi-source data. *International Journal of Applied Earth Observation and Geoinformation*, 2024, Vol. 128, art. 103714. DOI: 10.1016/j.jag.2024.103714 EDN: QKEYEO
14. Liu Y., Liu Q., Li X. A deep learning-based model for cold anticyclonic eddies and warm cyclonic eddies detection in the Kuroshio extension. *2022 Photonics & Electromagnetics Research Symposium (PIERS)*. Hangzhou, China, 2022. P. 258–263. DOI: 10.1109/PIERS55526.2022.9792954
15. Xu G., Xie W., Dong C., Gao X. Application of three deep learning schemes into oceanic eddy detection. *Frontiers in Marine Science*, 2021, Vol. 8, art. 672334. DOI: <https://doi.org/10.3389/fmars.2021.672334> EDN: LSUWBj
16. Gao X., Zhang W., Chen R. Mesoscale eddy identification in the North Pacific based on DeepLabV3+. *Journal of Physics: Conference Series*, 2025, Vol. 3007, № 1, art. 012052. DOI: 10.1088/1742-6596/3007/1/012052 EDN: QFMTLQ
17. Hou M., Fang L., Wu K., Yang J., Chen G. Multi-scale eddy identification and analysis based on deep learning method and ocean color data. *International Journal of Digital Earth*, 2025, Vol. 18, № 1, art. 2505624. DOI: <https://doi.org/10.1080/17538947.2025.2505624> EDN: WSTHFT
18. Liu Y., Gao L., Liu Q., Li X. Dual-branch neural network for mesoscale eddy identification based on multi-variables remote sensing data. *2023 Photonics & Electromagnetics Research Symposium (PIERS)*. Prague, Czech Republic, 2023. P. 275–279. DOI: 10.1109/PIERS59004.2023.10221436
19. Sun H., Li H., Xu M., Xia T., Han Z., Liu C. A data-fused lightweight CNN-Transformer model for ocean eddy detection. *Science of Remote Sensing*, 2026, Vol. 13, art. 100412. DOI: <https://doi.org/10.1016/j.srs.2026.100412>
20. Xie H., Xu Q., Dong C. Deep learning for mesoscale eddy detection with feature fusion of multisatellite observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, Vol. 17, P. 18351–18364. DOI: 10.1109/jstars.2024.3468457 EDN: NXKZZS
21. Li B., Tang H., Ma D., Lin J. A dual-attention mechanism deep learning network for mesoscale eddy detection by mining spatiotemporal characteristics. *Journal of Atmospheric and Oceanic Technology*, 2022, Vol. 39, Iss. 8, P. 1115–1128. DOI: <https://doi.org/10.1175/JTECH-D-21-0128.1> EDN: CUVQCO
22. Sun H., Li H., Xu M., Xia T., Yu H. Detecting ocean eddies with a lightweight and efficient convolutional network. *Remote Sensing*, 2024, Vol. 16, Iss. 24, art. 4808. DOI: <https://doi.org/10.3390/rs16244808> EDN: AJFWWY
23. Zhao N., Huang B., Zhang X., Ge L., Ge Chen. Intelligent identification of oceanic eddies in remote sensing data via Dual-Pyramid UNet. *Atmospheric and Oceanic Science Letters*, 2023, Vol. 16, Iss. 4, art. 100335. DOI: <https://doi.org/10.1016/j.aosl.2023.100335> EDN: LNAHY
24. Liu C., Lin X., Xu G., Han G., Liu Y. Improved identification and tracking of three-dimensional eddies in the Southern Ocean utilizing 3D-U-Res-Net. *Frontiers in Marine Science*, 2024, Vol. 11, art. 1482804. DOI: <https://doi.org/10.3389/fmars.2024.1482804> EDN: VBSFJC
25. Xu G., Xie W., Lin X., Liu Y., Hang R., Sun W., Liu D., Dong C. Detection of three-dimensional structures of oceanic eddies using artificial intelligence. *Ocean Modelling*, 2024, Vol. 190, art. 102385. DOI: <https://doi.org/10.1016/j.ocemod.2024.102385> EDN: MXSAHF
26. Liu X., Wang M. Detection of ocean eddies from satellite ocean color and SST measurements using a deep learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 2025, Vol. 144, art. 104929. DOI: 10.1016/j.jag.2025.104929 EDN: FBVPBK
27. Dong Ch., You Zh., Dong J., Ji J., Sun W., Xu G., Lu X., Xie H., Teng F., Liu Yu., Xu A., Wang Q., Xia Q., Lin X., Fu M., Wang J., Cao Yu., Han G. Oceanic mesoscale eddies. *Ocean-Land-Atmosphere Research*, 2025, vol. 4, art. 0081. DOI: 10.34133/olar.0081 EDN: OILJRV
28. Bearman A., Russakovsky O., Ferrari V., Fei-Fei L. What's the point: semantic segmentation with point supervision. *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*. Eds. B. Leibe, J. Matas, N. Sebe, M. Welling. Cham: Springer, 2016. Vol. 9911, P. 549–565. DOI: https://doi.org/10.1007/978-3-319-46478-7_34
29. Chelton D.B., Schlax M.G., Samelson R.M. Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, 2011, Vol. 91, Iss. 2, P. 167–216. DOI: <https://doi.org/10.1016/j.pcean.2011.01.002> EDN: YBYDIJ
30. Xie E., Wang W., Yu Z., Anandkumar A., Alvarez J.M., Luo P. SegFormer: simple and efficient design for semantic segmentation with transformers. *arXiv:2105.15203 [cs.CV]*. 2021. DOI: <https://doi.org/10.48550/arXiv.2105.15203>
31. Guangjun X., Yucheng S., Yang Y., Huarong X., Wenhong X., Jingyuan L., Xiayan L., Yu L., Changming D. Recent developments in ai-based oceanic eddy identification. *Journal of Marine Sciences*, 2024, Vol. 42, № 3, P. 38–50.

Информация об авторах

Игорь Александрович Ботыгин, кандидат технических наук, доцент отделения информационных технологий Инженерной школы информационных технологий и робототехники, Национальный исследовательский Томский политехнический университет, Российская Федерация, 634050, г. Томск, пр. Ленина, 30; bia@tpu.ru; <https://orcid.org/0000-0001-7720-734X>

Владимир Алексеевич Крутиков, доктор физико-математических наук, руководитель лаборатории геоинформационных технологий, Институт мониторинга климатических и экологических систем СО РАН, Российская Федерация, 634055, г. Томск, пр. Академический, 10/3; urmant@mail.ru; <https://orcid.org/0000-0002-6296-8151>.

Вероника Владимировна Кулак, студент отделения информационных технологий Инженерной школы информационных технологий и робототехники, Национальный исследовательский Томский политехнический университет, Российская Федерация, 634050, г. Томск, пр. Ленина, 30; vvk156@tpu.ru; <https://orcid.org/0009-0001-4869-4246>

Владислав Станиславович Шерстнев, кандидат технических наук, доцент отделения информационных технологий Инженерной школы информационных технологий и робототехники, Национальный исследовательский Томский политехнический университет, Российская Федерация, 634050, г. Томск, пр. Ленина, 30; vss@tpu.ru; <https://orcid.org/0000-0001-7230-575X>

Анна Игоревна Шерстнева, кандидат физико-математических наук, доцент отделения математики и математической физики Инженерной школы ядерных технологий, Национальный исследовательский Томский политехнический университет, Российская Федерация, 634050, г. Томск, пр. Ленина, 30; sherstneva@tpu.ru; <https://orcid.org/0000-0003-3326-825X>

Поступила в редакцию: 20.05.2026

Поступила после рецензирования: 26.06.2026

Принята к публикации: 30.06.2026

REFERENCES

1. Rosenfeld A., Pfaltz J.L. Sequential operations in digital picture processing. *Journal of the ACM*, 1966, vol. 13, Iss. 4, pp. 471–494. DOI: <https://doi.org/10.1145/321356.321357>
2. Zucker S.W. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 1976, vol. 5, Iss. 3, pp. 382–399. DOI: [https://doi.org/10.1016/S0146-664X\(76\)80014-7](https://doi.org/10.1016/S0146-664X(76)80014-7)
3. Otsu N. A Threshold selection method from gray-level histograms. *IEEE Transaction on Systems, Man and Cybernetics*, 1979, vol. 9, Iss. 1, pp. 62–66. DOI: 10.1109/TSMC.1979.4310076
4. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA, 2015. pp. 3431–3440. DOI: 10.1109/CVPR.2015.7298965
5. Kalinin M.A., Verezhenskaya P.S., Krinitskiy M.A., Borisov M.A., Tilinina N.D. Bayesian optimization of the method of identification of mesoscale ocean eddies in the Labrador sea in eddy-resolving modelling data. *Journal of oceanological research*, 2024, vol. 52, no. 4, pp. 56–73. (In Russ.) DOI: 10.29006/1564-2291.JOR-2024.52(4).4 EDN: FSRROH
6. Khachatryan E., Sandalyuk N., Loizou P. Eddy detection in the marginal ice zone with Sentinel-1 data using YOLOv5. *Remote Sensing*, 2023, vol. 15, Iss. 9, art. 2244. DOI: <https://doi.org/10.3390/rs15092244> EDN: GGCCEI
7. Kalinin M.A., Krinitskiy M.A., Verezhenskaya P.S. Detection of Irminger rings in high resolution ocean hydrodynamic modeling data using artificial neural networks. *Moscow University Physics Bulletin*, 2025, vol. 80, Suppl 3, pp. S1270–S1278. DOI: <https://doi.org/10.3103/S0027134925703102> EDN: PMSWYV
8. Liu Y., Zheng Q., Li X. Detection and analysis of mesoscale eddies based on deep learning. *Artificial intelligence oceanography*. Eds. X. Li, F. Wang. Singapore, Springer, 2023. pp. 209–225. DOI: https://doi.org/10.1007/978-981-19-6375-9_10.
9. Liu Y., Li X., Ren Y. A deep learning model for oceanic mesoscale eddy detection based on multi-source remote sensing imagery. *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, pp. 6762–6765. DOI: 10.1109/IGARSS39084.2020.9323716
10. Geng J., Gao H., Huang B., Radenkovic M., Chen G. ARU²-net: a deep learning approach for global-scale oceanic eddy detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, vol. 17, pp. 11997–12007. DOI: 10.1109/jstars.2024.3419175 EDN: AQBZOB
11. Sun X., Zhang M., Dong J., Lguensat R., Yang Y., Lu X. A deep framework for eddy detection and tracking from satellite sea surface height data. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, vol. 59, no. 9, pp. 7224–7234. DOI: 10.1109/TGRS.2020.3032523 EDN: BZLKWB
12. Zhao Y., Fan Z., Li H., Zhang R., Xiang W., Wang S., Zhong G. SymmetricNet: end-to-end mesoscale eddy detection with multi-modal data fusion. *Frontiers in Marine Science*, 2023, vol. 10, art. 1174818. DOI: <https://doi.org/10.3389/fmars.2023.1174818> EDN: ERMINB
13. Huo J., Zhang J., Yang Ju., Li Ch., Liu G., Cui W. High kinetic energy mesoscale eddy identification based on multi-task learning and multi-source data. *International Journal of Applied Earth Observation and Geoinformation*, 2024, vol. 128, art. 103714. DOI: 10.1016/j.jag.2024.103714 EDN: QKEYEO
14. Liu Y., Liu Q., Li X. A deep learning-based model for cold anticyclonic eddies and warm cyclonic eddies detection in the Kuroshio extension. *2022 Photonics & Electromagnetics Research Symposium (PIERS)*. Hangzhou, China, 2022. pp. 258–263. DOI: 10.1109/PIERS55526.2022.9792954

15. Xu G., Xie W., Dong C., Gao X. Application of three deep learning schemes into oceanic eddy detection. *Frontiers in Marine Science*, 2021, vol. 8, art. 672334. DOI: <https://doi.org/10.3389/fmars.2021.672334> EDN: LSUWBJ
16. Gao X., Zhang W., Chen R. Mesoscale eddy identification in the North Pacific based on DeepLabV3+. *Journal of Physics: Conference Series*, 2025, vol. 3007, no. 1, art. 012052. DOI: [10.1088/1742-6596/3007/1/012052](https://doi.org/10.1088/1742-6596/3007/1/012052) EDN: QFMTLQ
17. Hou M., Fang L., Wu K., Yang J., Chen G. Multi-scale eddy identification and analysis based on deep learning method and ocean color data. *International Journal of Digital Earth*, 2025, vol. 18, no. 1, art. 2505624. DOI: <https://doi.org/10.1080/17538947.2025.2505624> EDN: WSTHFT
18. Liu Y., Gao L., Liu Q., Li X. Dual-branch neural network for mesoscale eddy identification based on multi-variables remote sensing data. *2023 Photonics & Electromagnetics Research Symposium (PIERS)*. Prague, Czech Republic, 2023. pp. 275–279. DOI: [10.1109/PIERS59004.2023.10221436](https://doi.org/10.1109/PIERS59004.2023.10221436)
19. Sun H., Li H., Xu M., Xia T., Han Z., Liu C. A data-fused lightweight CNN-Transformer model for ocean eddy detection. *Science of Remote Sensing*, 2026, vol. 13, art. 100412. DOI: <https://doi.org/10.1016/j.srs.2026.100412>
20. Xie H., Xu Q., Dong C. Deep learning for mesoscale eddy detection with feature fusion of multisatellite observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, vol. 17, pp. 18351–18364. DOI: [10.1109/jstars.2024.3468457](https://doi.org/10.1109/jstars.2024.3468457) EDN: NXKZZS
21. Li B., Tang H., Ma D., Lin J. A dual-attention mechanism deep learning network for mesoscale eddy detection by mining spatiotemporal characteristics. *Journal of Atmospheric and Oceanic Technology*, 2022, vol. 39, Iss. 8, pp. 1115–1128. DOI: <https://doi.org/10.1175/JTECH-D-21-0128.1> EDN: CUVQCO
22. Sun H., Li H., Xu M., Xia T., Yu H. Detecting Ocean Eddies with a Lightweight and Efficient Convolutional Network. *Remote Sensing*, 2024, vol. 16, Iss. 24, art. 4808. DOI: <https://doi.org/10.3390/rs16244808> EDN: AJFWWY
23. Zhao N., Huang B., Zhang X., Ge L., Ge Ch. Intelligent identification of oceanic eddies in remote sensing data via Dual-Pyramid UNet. *Atmospheric and Oceanic Science Letters*, 2023, vol. 16, Iss. 4, art. 100335. DOI: <https://doi.org/10.1016/j.aosl.2023.100335> EDN: LNARHY
24. Liu C., Lin X., Xu G., Han G., Liu Y. Improved identification and tracking of three-dimensional eddies in the Southern Ocean utilizing 3D-U-Res-Net. *Frontiers in Marine Science*, 2024, vol. 11, art. 1482804. DOI: <https://doi.org/10.3389/fmars.2024.1482804> EDN: VBSFJC
25. Xu G., Xie W., Lin X., Liu Y., Hang R., Sun W., Liu D., Dong C. Detection of three-dimensional structures of oceanic eddies using artificial intelligence. *Ocean Modelling*, 2024, vol. 190, art. 102385. DOI: <https://doi.org/10.1016/j.ocemod.2024.102385> EDN: MXSAHF
26. Liu X., Wang M. Detection of ocean eddies from satellite ocean color and SST measurements using a deep learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 2025, vol. 144, art. 104929. DOI: [10.1016/j.jag.2025.104929](https://doi.org/10.1016/j.jag.2025.104929) EDN: FBVPBK
27. Dong Ch., You Zh., Dong J., Ji J., Sun W., Xu G., Lu X., Xie H., Teng F., Liu Yu., Xu A., Wang Q., Xia Q., Lin X., Fu M., Wang J., Cao Yu., Han G. Oceanic mesoscale eddies. *Ocean-Land-Atmosphere Research*, 2025, vol. 4, art. 0081. DOI: [10.34133/olar.0081](https://doi.org/10.34133/olar.0081) EDN: OILJRV
28. Bearman A., Russakovsky O., Ferrari V., Fei-Fei L. What's the point: semantic segmentation with point supervision. *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*. Eds. B. Leibe, J. Matas, N. Sebe, M. Welling. Cham, Springer, 2016. Vol. 9911, pp. 549–565. DOI: https://doi.org/10.1007/978-3-319-46478-7_34
29. Chelton D.B., Schlax M.G., Samelson R.M. Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, 2011, vol. 91, Iss. 2, pp. 167–216. DOI: <https://doi.org/10.1016/j.pocean.2011.01.002> EDN: YBYDIJ
30. Xie E., Wang W., Yu Z., Anandkumar A., Alvarez J.M., Luo P. SegFormer: simple and efficient design for semantic segmentation with transformers. *arXiv:2105.15203 [cs.CV]*. 2021. DOI: <https://doi.org/10.48550/arXiv.2105.15203>
31. Guangjun X., Yucheng S., Yang Y., Huarong X., Wenhong X., Jingyuan L., Xiayan L., Yu L., Changming D. Recent developments in ai-based oceanic eddy identification. *Journal of Marine Sciences*, 2024, vol. 42, no. 3, pp. 38–50.

Information about the authors

Igor A. Botygin, Cand. Sci. (Eng.), Associate Professor, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation; bia@tpu.ru; <https://orcid.org/0000-0001-7720-734X>

Vladimir A. Krutikov, Dr. Sci. (Phys.-Math.), Head of the Laboratory, Institute of Monitoring of Climatic and Ecological Systems of the SB RAS, 10/3, Akademicheskyy avenue, Tomsk, 634055, Russian Federation; urmant@mail.ru; <https://orcid.org/0000-0002-6296-8151>.

Veronika V. Kulak, Student, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation; vvk156@tpu.ru; <https://orcid.org/0009-0001-4869-4246>

Vladislav S. Sherstnev, Cand. Sci. (Eng.), Associate Professor, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation; vss@tpu.ru; <https://orcid.org/0000-0001-7230-575X>

Anna I. Sherstneva, Cand. Sci. (Phys.-Math.), Associate Professor, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation; sherstneva@tpu.ru; <https://orcid.org/0000-0003-3326-825X>

Received: 20.05.2026

Revised: 26.06.2026

Accepted: 30.06.2026