

Автоматизация обнаружения нерки в данных визуальной оптической съемки с борта БПЛА в районе озера Начикинское (Камчатка) с применением искусственных нейронных сетей

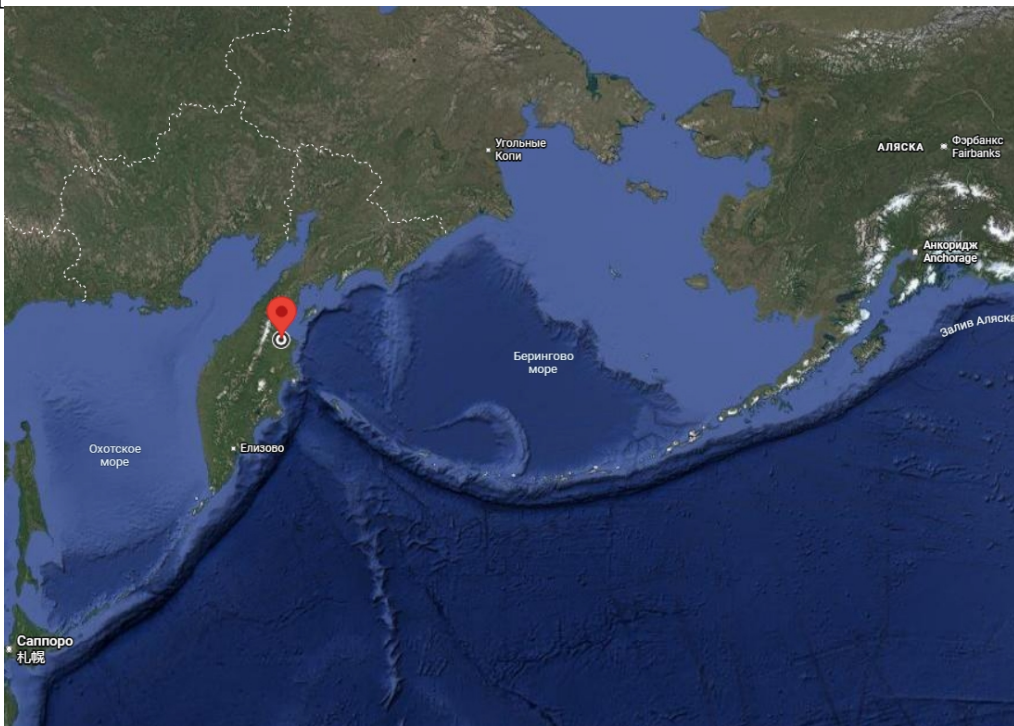
Шаталов Дмитрий¹, Криницкий Михаил^{1,2}, Шубкин Сергей³

(1) МФТИ, (2) ИО РАН, (3) Камчатский филиал ГНЦ РФ ФГБНУ «ВНИРО» («КамчатНИРО»).

Актуальность

Биологическая и хозяйственная значимость

Нерка (*Oncorhynchus nerka*) — ключевой вид тихоокеанской лосося для экосистем и экономики Камчатки. Озеро Начикинское — одно из важнейших нерестилищ. Мониторинг необходим для оценки популяций, прогнозирования запасов и рационального промысла.



Проблема обработки данных

БПЛА-съёмка даёт ортофотопланы сверхвысокого разрешения, но их ручная обработка крайне трудоёмка, субъективна и не позволяет оперативно получать результаты. Это создаёт запрос на автоматизацию учёта.



Цель

Разработка и апробация метода автоматического обнаружения особей нерки на ортофотопланах с применением нейросетевых моделей семейства YOLO как одного из перспективных подходов для последующего анализа поведения.

План работы

1. Обработать данные, собранные в виде ортофотоплана на оз. Начинское с помощью БПЛА «Геоскан-201» с экспертной разметкой нерки.
2. Обучить модели семейства YOLO на получившемся в результате обработки датасете.
3. Протестировать обученные модели на части датасета, провести сравнительный анализ по выбранным метрикам.

Данные и методы

Исходные данные:

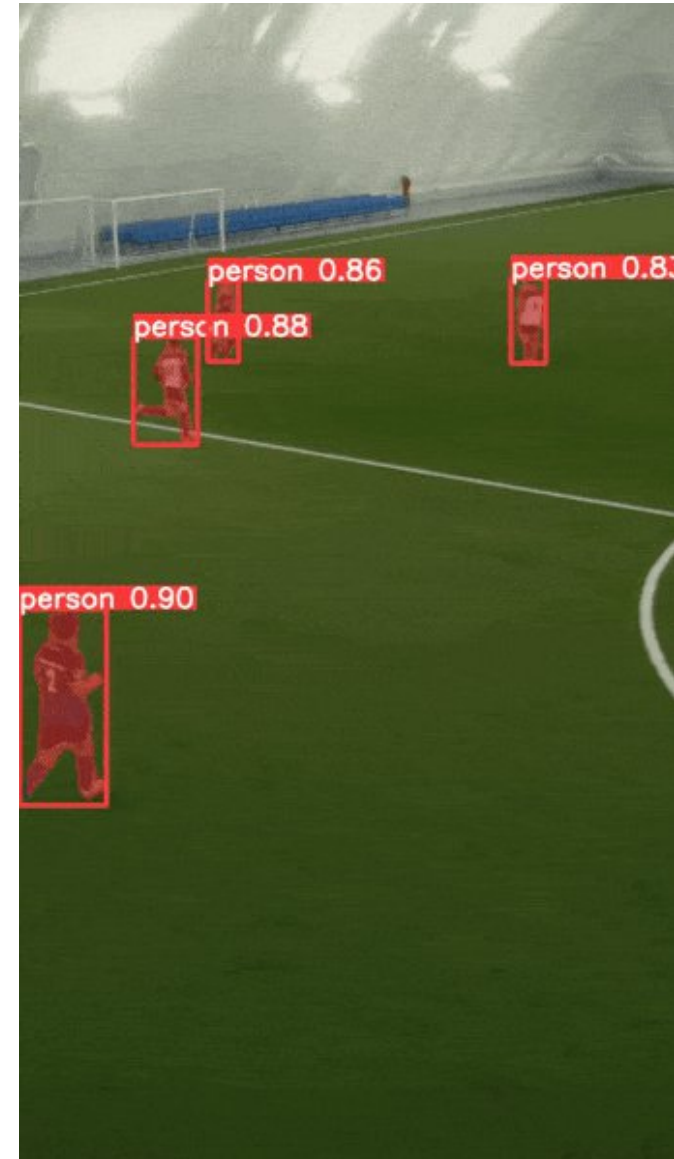
- Аэрофотосъёмка побережья и мелководных участков нерестилищ оз. **Начикинское**
- Съёмка осуществлялась с помощью БПЛА самолетного типа «**Геоскан 201М**»
- Построен ортофотоплан высокого разрешения в программе **Agisoft Metashape Professional**
- **Разрешение съёмки**, см/пик: 1,13



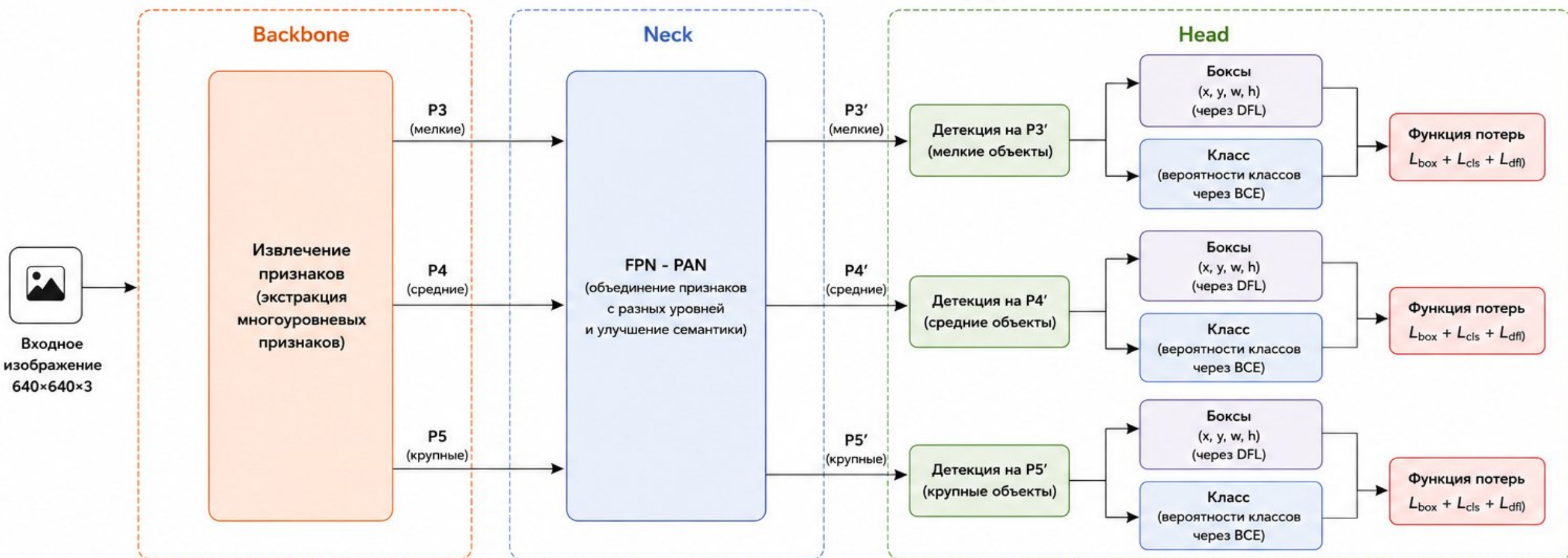
Данные и методы

Почему YOLO(YouOnlyLookOnce) ?

- **Высокая скорость обработки** — алгоритм работает в режиме одного прохода (single-stage)
- **Готовая экосистема** — единый и популярный фреймворк для детекции и оценки позы.
- **Предобученные модели** - инициализированы весами, обученными на датасете COCO (118 тыс. изображений, 80 классов).



YOLOv8 – Архитектура



Расшифровка аббревиатур:

FPN (Feature Pyramid Network)

Пирамидальная сеть признаков. Объединяет признаки разных уровней для лучшего представления объектов разного масштаба.

PAN (Path Aggregation Network)

Сеть агрегации путей. Улучшает поток информации между уровнями признаков, усиливая локализацию и семантику.

DFL (Distribution Focal Loss)

Функция потерь для регрессии боксов. Моделирует координаты как распределение вероятностей, что повышает точность локализации.

BCE (Binary Cross-Entropy)

Функция потерь для классификации. Оценивает разницу между предсказанными вероятностями классов и истинными метками.

Данные и методы

Общая функция потерь:

$$\mathcal{L}_{total} = \underbrace{7.5 \cdot \mathcal{L}_{box} + 0.5 \cdot \mathcal{L}_{cls} + 1.5 \cdot \mathcal{L}_{dfl}}_{\text{YOLOv8 (детекция)}} + \underbrace{\lambda_{pose} \cdot \mathcal{L}_{pose}}_{\text{Отдельно в YOLOv8-pose}}$$

2) $\mathcal{L}_{cls}$ — потеря классификации (BCE Loss)

$$\mathcal{L}_{cls} = -\frac{1}{N} \sum_{i=1}^N \left[y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \right]$$

Отвечает за правильность предсказания класса объекта для каждой положительной ячейки/предсказания.

3) $\mathcal{L}_{dfl}$ — Distribution Focal Loss (DFL Loss)

$$\mathcal{L}_{dfl} = -\frac{1}{N} \sum_{i=1}^N \sum_{k=0}^{K-1} w_k^* \log(p_{ik})$$

где K — число бинов (каналов) для распределения координаты,
 p_{ik} — предсказанная вероятность k -го бина для i -й координаты,
 w_k^* — целевое распределение (линейная интерполяция):

Отвечает за точность распределения координат боксов через дискретизацию (улучшает регрессию с помощью распределений).

4) $\mathcal{L}_{pose}$ — потеря позы (только в YOLOv8-pose)

Пусть M — число ключевых точек.

Для каждой ключевой точки j :

$$\mathcal{L}_{kpt}^{(j)} = \delta_j \cdot \left(\|\hat{\mathbf{p}}_j - \mathbf{p}_j^{gt}\|_2^2 \right)$$

где $\hat{\mathbf{p}}_j = (\hat{x}_j, \hat{y}_j)$ — предсказанная координата,

$\mathbf{p}_j^{gt} = (x_j^{gt}, y_j^{gt})$ — истинная координата,

$\delta_j \in \{0, 1\}$ — видимость точки (1 — видима, 0 — нет).

Общая потеря позы:

$$\mathcal{L}_{pose} = \frac{1}{M_{vis}} \sum_{j=1}^M \mathcal{L}_{kpt}^{(j)}$$

где $M_{vis} = \sum_{j=1}^M \delta_j$ — число видимых ключевых точек.

Отвечает за точность предсказания координат ключевых точек (поза человека/объекта).

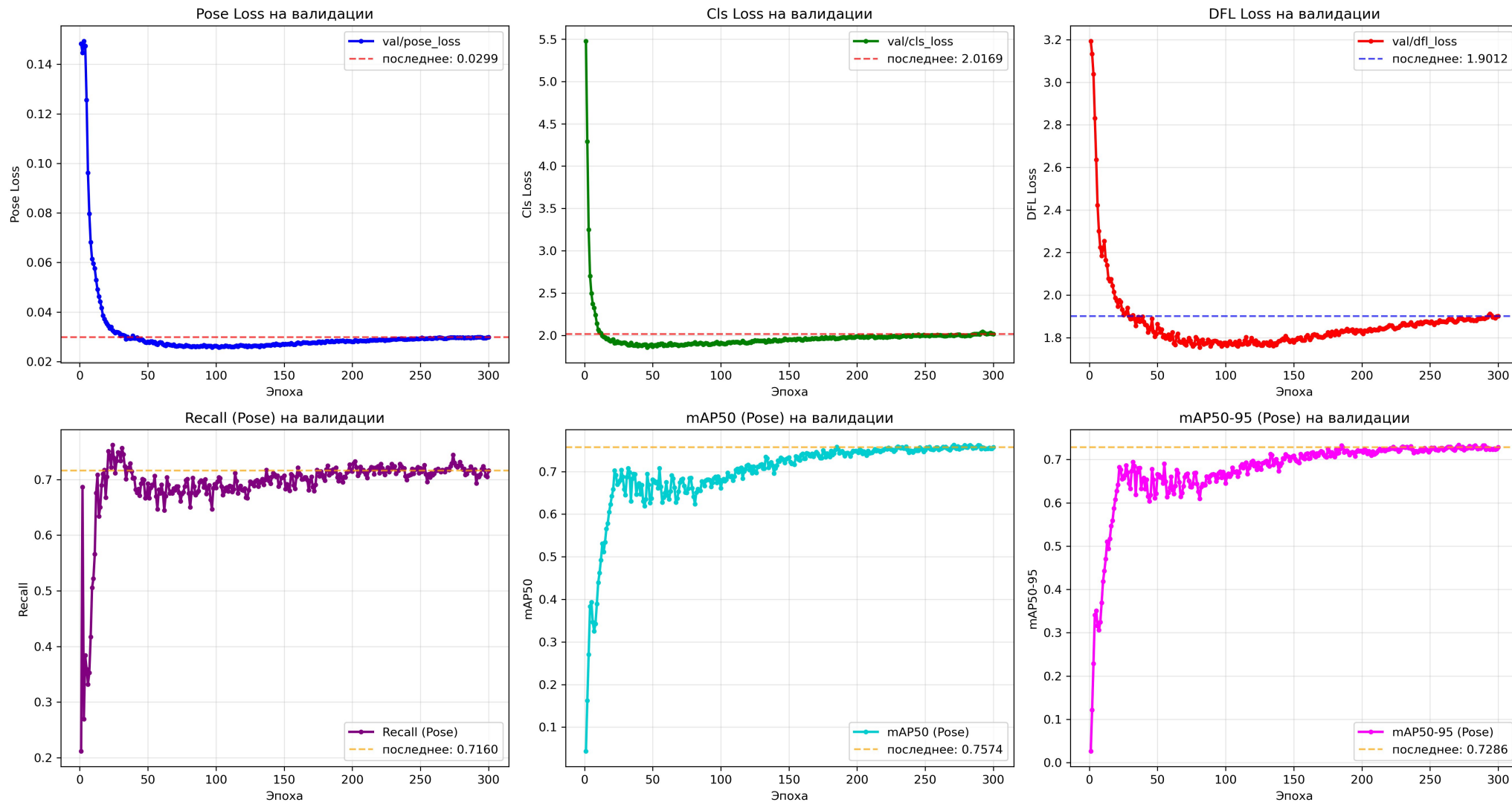
Обработка данных



Результаты

Обучение моделей

Динамика потерь и метрик на валидации (YOLOv8-Pose)



Результаты Тестирование моделей



Результаты

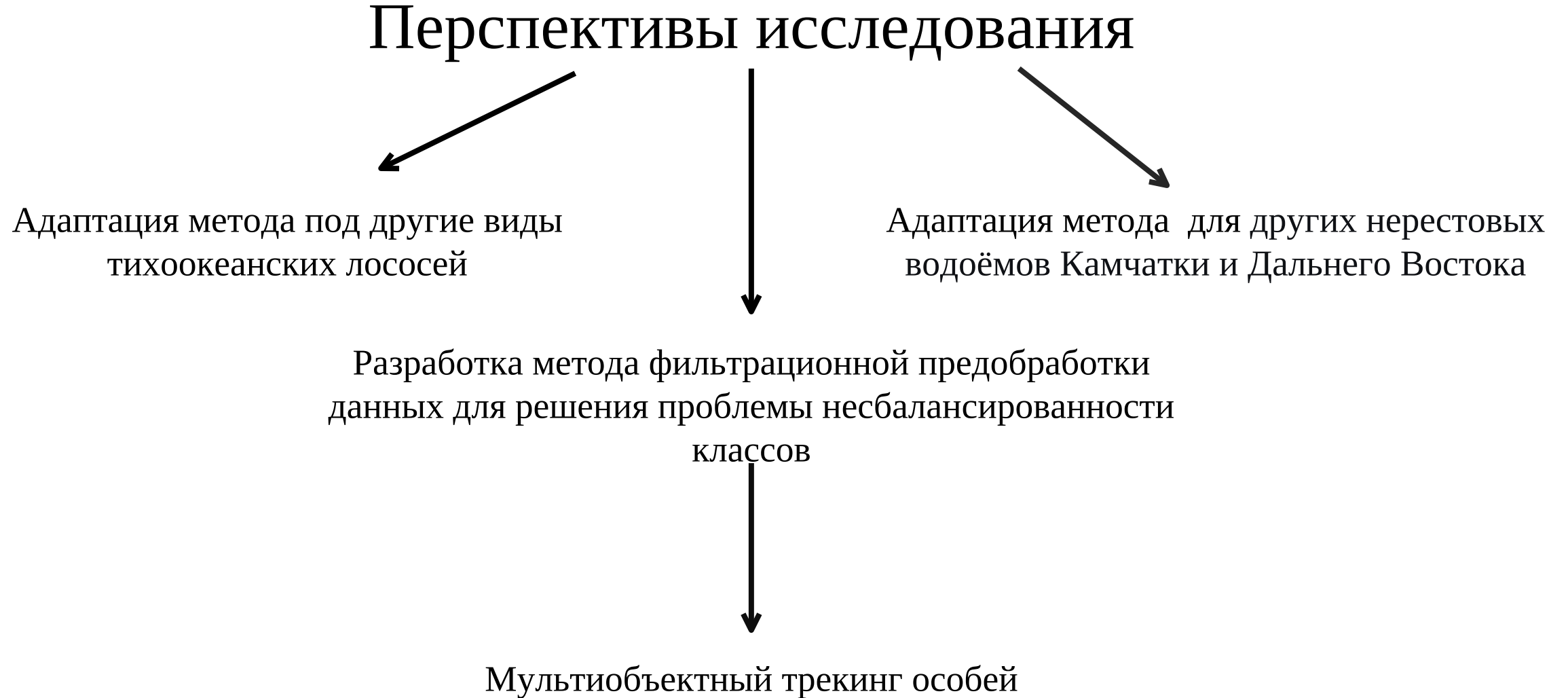
Тестирование моделей (confidence = 0.15)

Model	Type	Precision	Recall	F1	mAP50	Ratio
yolov8n	det	0.929	0.359	0.518	0.112	0.3877
yolov8s	det	0.926	0.345	0.503	0.159	0.372
yolov8m	det	0.957	0.333	0.494	0.159	0.348
yolov8l	det	0.873	0.541	0.668	0.200	0.619
yolo26n	det	0.932	0.039	0.074	0.005	0.042
yolo26m	det	0.973	0.070	0.131	0.014	0.072
yolov8n-pose	pose	0.923	0.439	0.595	0.289	0.475
yolov8s-pose	pose	0.845	0.630	0.722	0.436	0.745
yolov8m-pose	pose	0.853	0.655	0.741	0.506	0.768
yolov8l-pose	pose	0.751	0.856	0.800	0.642	1.140

Обсуждение и выводы

- 1. Метод автоматического обнаружения нерки на основе YOLO показал свою применимость.** Лучшая модель из исследуемых (YOLOv8l-pose) обеспечивает $F1=0.8$, и находит **более 85% нерки**, что подтверждает эффективность подхода.
- 2. Модели Yolo - pose показали преимущество перед детекционными.** Использование ключевых точек позволяет точнее локализовать местонахождение объекта и избежать проблем, связанных с информационным шумом от bounding boxes.
- 3. Основным ограничением точности являются множественные детекции одного объекта**, а не пропуски. Решение — оптимизация параметров постобработки (порог IoU).

Перспективы исследования



```
graph TD; A[Перспективы исследования] --> B[Адаптация метода под другие виды тихоокеанских лососей]; A --> C[Адаптация метода для других нерестовых водоёмов Камчатки и Дальнего Востока]; A --> D[Разработка метода фильтрационной предобработки данных для решения проблемы несбалансированности классов]; D --> E[Мультиобъектный трекинг особей];
```

Адаптация метода под другие виды
тихоокеанских лососей

Адаптация метода для других нерестовых
водоёмов Камчатки и Дальнего Востока

Разработка метода фильтрационной предобработки
данных для решения проблемы несбалансированности
классов

Мультиобъектный трекинг особей

Спасибо за внимание!!!