

Сравнение YOLO и RT-DETR для спутниковых изображений высокого разрешения

В данном документе приведено сравнение моделей YOLO и RT-DETR для задач обнаружения объектов на спутниковых изображениях, включая работу с изображениями сверхвысокого разрешения (до 50 000 × 50 000 пикселей).

1. Краткий обзор моделей

YOLO (You Only Look Once) — семейство сверточных моделей реального времени, ориентированных на высокую скорость инференса.

RT-DETR (Real-Time DEtection TRansformer) — трансформерная модель обнаружения объектов, использующая глобальное внимание и не требующая NMS.

2. Сравнение точности и характеристик

Модель	Архитектура	Точность (mAP, обобщенно)	Преимущества	Недостатки
YOLO (v5/v8)	CNN	Высокая на мелких и средних объектах	Скорость, простота, edge-устройства	Ограниченный глобальный контекст
RT-DETR	Transformer	Сравнима или выше на сложных сценах	Глобальное внимание, NMS-free	Высокие требования к GPU

3. Работа с изображениями сверхвысокого разрешения (50 000 × 50 000)

Большинство моделей не способны напрямую обрабатывать изображения такого размера из-за ограничений памяти. Поэтому применяются методы разбиения на тайлы (tiling).

4. Алгоритмы разбиения на тайлы

- 1. Классический Sliding Window (скользящее окно)
 - Изображение разбивается на перекрывающиеся тайлы (например, 1024×1024).
 - Перекрытие: 10–30% для предотвращения потери объектов на границах.
 - Подходит для YOLO и RT-DETR.
- 2. Adaptive Tiling (адаптивное тайлирование)
 - Размер тайла изменяется в зависимости от плотности объектов.

- Часто используется предварительный downscale + coarse detection.

3. Grid Tiling без перекрытия

- Быстрое разбиение на сетку без overlap.
- Риск потери объектов на границах.

4. Pyramid / Multi-scale Tiling

- Изображение анализируется на нескольких масштабах.
- Особенно эффективно для мелких объектов (самолёты, корабли).

5. Patch-based inference + merging

- Детекция на тайлах.
- Слияние результатов с помощью IoU-based merging или Soft-NMS.

5. Постобработка и объединение результатов

- Удаление дубликатов с помощью IoU-порогов.
- Global NMS или Soft-NMS.
- Пересчет координат bounding box в глобальную систему изображения.

6. Лицензии

Модель	Лицензия	Комментарии
YOLO (Ultralytics)	AGPL-3.0	Коммерческое использование требует отдельной лицензии
RT-DETR	Apache-2.0	Разрешено коммерческое использование

7. Выводы

Для спутниковых изображений сверхвысокого разрешения ключевым фактором является грамотное тайлирование. YOLO удобен для быстрого и масштабируемого инференса, в то время как RT-DETR перспективен для сложных сцен с большим количеством объектов.