

Трекинг игроков в спорте

мартовский чекпоинт

Обновление: новый датасет

- Я подключил к проекту новый датасет [SportsMOT](#)
- Датасет состоит из кадров из спортивных трансляций (футбол, баскетбол, волейбол), на которых размечены bbox-ы игроков
- Датасет содержит как подмножество предыдущий датасет, поэтому с этого момента пользуюсь только им
- Я сконвертировал датасет в формат YOLO

Технические ограничения

- Эксперименты на новом датасете занимают достаточно долго
- Поэтому я поставил следующие ограничения для экспериментов:
 - Обучаюсь только на 20% датасета
 - а именно, взял первые 20% кадров из каждого клипа в датасете
 - Ограничиваю эксперименты ≈ 2 часами обучения на GPU P100
- Лучшие модели по результатам экспериментов обучаю на всём датасете без ограничения по времени

Эксперименты: подбор learning rate на маленькой модели

модель	млн параметров	число эпох	начальный lr	конечный lr как доля от начального	число warmup эпох	лучший mAP@50-95 на валидации
yolo11s	9.4	20	4e-3	0.01	3	0.786
			2e-3	0.5	3	0.775
			2e-3	0.01	0	0.761
			2e-3	0.01	3	0.792
			1.5e-3	0.01	3	0.796
			1e-3	0.01	3	0.800
			5e-4	0.01	3	0.800

Эксперименты: размер YOLO модели

модель	млн параметров	число эпох	начальный lr	конечный lr как доля от начального	число warmup эпох	лучший mAP@50-95 на валидации
yolo11n	2.6	20	2e-3	0.01	3	0.740
yolo11s	9.4	20	2e-3			0.792
			5e-4			0.800
yolo11m	20.1	14	2e-3			0.815
			5e-4			0.821
yolo11l	25.3	10	5e-4			0.825

Эксперименты: архитектура RT-DETR

- Я попробовал архитектуру, отличную от YOLO, но пока что она даёт хуже результаты

модель	млн параметров	число эпох	начальный lr	конечный lr как доля от начального	число warmup эпох	лучший mAP@50-95 на валидации
rtdetr-l	32.8	6	2e-3	0.01	3	0.767
			5e-4		3	0.794
			5e-4		1	0.797

Итоги

- При обучении на 100% данных лучшая модель (yolo11l, lr=5e-4) набирает $mAP@50-95 = \mathbf{0.855}$ на валидационной выборке
- Это побивает обученный на старом датасете baseline ($mAP@50-95 = \mathbf{0.466}$)
- Это также побивает модель, аналогичную baseline, но обученную на новом датасете (сравнение при обучении на 20% данных: $mAP@50-95 = \mathbf{0.792}$ vs $\mathbf{0.825}$)
 - т.е. улучшение состоит не только в обновлении датасета