

## **9. Генерация данных**

Ваша команда столкнулась с недостатком данных для обучения моделей детектирования. Было принято решение рассмотреть возможность синтетического расширения обучающего набора. Ваша задача состоит в том, что бы проанализировать различные методы генерации данных для задачи детектирования и предоставить заключение о целесообразности их использования и необходимых ресурсах.

### **Минимальные требования:**

- Выбран датасет для детектирования, на основе которого будет проводится оценка эффективности.
- Исследовано не менее 2 уникальных подходов (например, GAN, Diffusions и т. д.)
- Каждым из подходов получен синтетический датасет, смешан с исходным (обучающим) и обучена YOLO
- Проведено сравнение тестовых метрик (реальные данные исходного датасета) для каждого из подходов и «чистого» (исходный трейн) обучения
- Написан аналитический отчет (например, в jupyter-тетрадке) с выводами о целесообразности применения генерации данных

### **Дополнительные улучшения:**

- Дополнительные подходы к улучшению (более 2 из минимальных требований)

**Время выполнения: 2—3 месяца**